

83186

Access DB# \_\_\_\_\_

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: JANE ZARA Examiner #: 77512 Date: 12/30/02  
 Art Unit: 1635 Phone Number 306-5820 Serial Number: 09/915,814  
 Mail Box and Bldg/Room Location: 11D03 Results Format Preferred (circle): PAPER DISK E-MAIL  
11E12

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: AS to H-SL ExpressionInventors (please provide full names): Burton et alEarliest Priority Filing Date: 7/26/01

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please Search Seq ID #3

- size unlimited

- size limit to 100 NT's.

Thanks

RECEIVED  
 DEC 30 2002  
 STIC

NA 3804

Edward Hart  
 Technical Info. Specialist  
 STIC/Biotech  
 CMI 6B02 Tel: 305-9203

## STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>2</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>12/03</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>1/3/03</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>06</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 03:04:20 ; Search time 6400 Seconds  
(without alignments)  
17297.980 Million cell updates/sec

Title: US-09-915-814-3  
Perfect score: 3804  
Sequence: 1 ctcttgtaagagagtgcta.....taataaagtatttaatra 3804

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2054640 segs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

GenEmbl:  
1: gb\_da:  
2: gb\_hlg:  
3: gb\_in:  
4: gb\_cm:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_sts:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vl:  
15: em\_da:  
16: em\_fun:  
17: em\_hum:  
18: em\_in:  
19: em\_mu:  
20: em\_cm:  
21: em\_or:  
22: em\_ov:  
23: em\_pat:  
24: em\_ph:  
25: em\_pl:  
26: em\_ro:  
27: em\_sts:  
28: em\_un:  
29: em\_vl:  
30: em\_hlg\_hum:  
31: em\_hlg\_inv:  
32: em\_hlg\_other:  
33: em\_hlg\_mus:  
34: em\_hlg\_pln:  
35: em\_hlg\_rod:  
36: em\_hlg\_man:  
37: em\_hlg\_vit:  
38: em\_sy:  
39: em\_higo\_hum:  
40: em\_higo\_mus:  
41: em\_higo\_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3804	100.0	3804	6 AX113509	AX113509 Sequence
2	3804	100.0	3804	9 HSU40002	U40002 Human hormo
3	2650.4	69.7	3255	9 HUMSLA	L11706 Human hormo
4	1896.4	49.9	3704	10 RNU40001	U40001 Rattus norv
5	1841.8	48.4	2792	10 AF177401	AF177401 Spermophi
6	1796.6	47.2	3125	4 SSHSLIPAS	AJ000482 Sus scrof
7	1636.4	43.0	3226	10 RATLIP	X51415 Rattus sp.
8	1620	42.6	2678	10 BC021642	BC021642 Mus muscu
9	1619.2	42.6	3172	10 MMU08188	U08188 Mus musculu
10	1155.2	30.4	168586	9 AC011497	AC011497 Homo sapi
11	552.8	14.5	652	4 AF163330	AF163330 Canis fam
12	538.2	14.1	656	4 AF141958	AF141958 Sus scrof
13	531.6	14.0	140808	2 AC021857	AC021857 Homo sapi
14	497.2	13.1	177575	2 AC021857	AJ000483 Sus scrof
15	436	11.5	2222	4 SSHSL5UTR	AJ224692 Sus scrof
16	405.6	10.7	498	4 SRA14692	AF179427 Mus muscu
17	392.8	10.3	17835	10 AF179427	AF179427 Mus muscu
18	392.8	10.3	190860	2 AC073803	AC073803 Mus muscu
19	390.2	10.3	215234	2 AC073745	AC073745 Mus muscu
20	388.6	10.2	131063	2 AC130942	AC130942 Rattus no
21	388.6	10.2	155817	2 AC121639	AC121639 Rattus no
22	282.8	7.4	142394	2 AC109038	AJ006076 Sus scrof
23	272	7.2	4211	4 SSC6076	AC130942 Rattus no
24	269.6	7.1	131063	2 AC130942	G15903 human STS C
25	258	6.8	258	11 G15903	AJ237675 Bos tauru
26	256.6	6.7	311	4 BTA237675	AC121639 Rattus no
27	256.6	6.7	155817	2 AC121639	U78042 Bos tauru
28	253	6.7	429	4 BRU78042	AJ060675 Sus scrof
29	245.4	6.5	1513	4 SSC6076	AC004321 Drosophi
30	225.4	5.9	34918	3 AC004321	AC020205 Drosophi
31	225.4	5.9	118533	2 AC020205	AC099023 Drosophi
32	225.4	5.9	180814	3 AC099023	AC099012 Drosophi
33	225.4	5.9	183592	3 AC099012	AE003792 Drosophi
34	225.4	5.9	256764	3 AE003792	S80110 hormone-sen
35	186.6	4.9	225	4 S80110	AC073803 Mus muscu
36	171	4.5	190860	2 AC073803	AF051867 Drosophi
37	149.8	3.9	2328	3 AY051867	AC109038 Rattus no
38	112.8	3.0	142394	2 AC109038	D90903 Synechocyst
39	102	2.7	106209	1 D90903	M17877 P.falciparu
40	99.8	2.6	2359	3 PFARAA	AC0663 P.falciparu
41	96.6	2.5	3187	6 A00663	AC113928 Rattus no
42	96.4	2.5	137721	2 AC113928	Y17332 Zea mays MR
43	94.6	2.5	1471	8 ZMPROPR	AJ130830 Zea mays
44	94.6	2.5	1792	8 ZMA130830	AY061106 Drosophi
45	87.4	2.3	4186	3 AY061106	

## ALIGNMENTS

RESULT 1	AX113509	3804 bp	mrna	linear	PAT 01-MAY-2001
LOCUS	AX113509				
DEFINITION	Sequence 1 from Patent WO0126664.				
ACCESSION	AX113509				
VERSION	AX113509.1				
KEYWORDS	GI:13939727				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 3804)				
JOURNAL	Mitchell, G.A. and Wang, S.P.				
	Hormone-sensitive lipase activity mediated male infertility				
	Patent: WO 0126664-A 1 19-APR-2001;				

**THIS PAGE BLANK (USPTO)**

FEATURES  
Source Hospital Sainte-Justine (CA)  
Location/Qualifiers  
1. 3804  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
278. 3508  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="Cac37855.1"  
/db\_xref="gi:13939728"  
translation="MEPSKSVSRSDMOPEHOPRIPLERPEKPTIAPOESKTIQ  
SNTQKPSANORPLTOETPAOHDAESOKERAPKASOEPRIAPKAPQSPYIT  
RVLITQDEASQSPGICKESTIQEERALRHRVAVQRPGEPPPAQDEESTPAAQ  
AKFGAKSEVPAPLESTQETPEOSDKOTPVQAKSKOGSTELGFTLKQELSTQRS  
ALRWKALSEWVADSESDVSSDTPATMGVMVQVSGVYKSGPYVMGYSQ  
TSPHETRSARHNYOPASRLHMDLRTQSLVLAELNIFESSQGEFRAQL  
SVFAGREOALGLEPALGRLLVHLDEPRLPANGRSLENTVIRARCLLHHSR  
YVANSRSTFERTSHNLAEIAYLATOLRAIYVYVORLITVRPGLPEEGEGIT  
ADPLREYVITLHKGCFYRCIGFOTPAIRPLOTISTGLVSFGEHYKRNELSVAS  
SLTSGRFATIDPELGAERITQNLVDHFKAWNTTEMEVLSLAMASATVRSR  
LSLPEAPFEMPLTADPTLVITISPLAHGPGVPLVRLISYDREGDSELSLIK  
SNGORSLTELPPOQAPRSRLIVHFGGGEVQVQTSHEBYLKSMAOELAPITSID  
YSLAPEAPFRALEECPFAVCMAIKHCLIGSTGERICLAGDSGACFVALRAAA  
YGVAPDGIAMAAYPATMLOPASPSRLISMDPLISVSKYSATNAGATBDHNS  
DOKALSMGLVRDRLALLDFRLGASSWLSLELSGRKSQKSEPIAEPMRSVSE  
AALAQPGPLGTDLSLKNLTLRLDLSLGNSETSDTPMSLSAETLSPTSDVPLFLP  
PEDAGEAEAKNELSPMDRGLGVRAAPFEGHPRRSOGATVPLSPYKNEFMS  
LIAPSMLKSLPVHIVACALDPLDSDVMILRLNIGCPVTLRAVEDLPHGLTLIA  
ALCREFROAELEVERILVLTLPAGAPSEGTGAAGVDGCGGRH"  
BASE COUNT 828 a 1225 c 1091 g 660 t  
ORIGIN  
Query Match 100.0%; Score 3804; DB 6; Length 3804;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCTTTAAGAGAGTGTAGGCACATAGCCCCCTCTATTCTTATCTCCACCAAG 60  
Db 1 CTTCTTTAAGAGAGTGTAGGCACATAGCCCCCTCTATTCTTATCTCCACCAAG 60  
QY 61 AAGAGGACAGAGTCTTACTTAGTGGGGCCAGCTGTGATGGCCAACTGCCAGCTG 120  
Db 61 AAGAGGACAGAGTCTTACTTAGTGGGGCCAGCTGTGATGGCCAACTGCCAGCTG 120  
QY 121 CTTAAAAAGAGAGACAGATGATGCTAGATGAGTGAACCCCAAGAGAACTGCCATCA 180  
Db 121 CTTAAAAAGAGAGACAGATGATGCTAGATGAGTGAACCCCAAGAGAACTGCCATCA 180  
QY 181 TGAGGAATCATGAGATCTGTGAAGAGAGAGGCTGGTGGGAGCCCAAGAGATAGA 240  
Db 181 TGAGGAATCATGAGATCTGTGAAGAGAGAGGCTGGTGGGAGCCCAAGAGATAGA 240  
QY 241 ACCTGAAGATCATATCTCCGTGAGGGAATAACATGAGAGCCAGGTTCTAAATCAGT 300  
Db 241 ACCTGAAGATCATATCTCCGTGAGGGAATAACATGAGAGCCAGGTTCTAAATCAGT 300  
QY 301 GTCTAGGTACAGTGGCACTGAAACAACACAGAGGCTATTAACCCGCTAGAGCCCTGG 360  
Db 301 GTCTAGGTACAGTGGCACTGAAACAACACAGAGGCTATTAACCCGCTAGAGCCCTGG 360  
QY 361 GCCGAAAAGACACCCATAGGCCAGCAAGATCGAAGACTCGCAGGAGTCCAAATACCA 420  
Db 361 GCCGAAAAGACACCCATAGGCCAGCAAGATCGAAGACTCGCAGGAGTCCAAATACCA 420  
QY 421 ACAGAACTCTCTCAAAACCAAGACCCCTCACCAGAGAGAGACCCCTGCACAACATGA 480  
Db 421 ACAGAACTCTCTCAAAACCAAGACCCCTCACCAGAGAGAGACCCCTGCACAACATGA 480  
QY 481 TGTGTAATCCAGAGAACTAGAGCCCAACAAAATCTGCTTCAAGAGAAATTTCT 540  
Db 481 TGTGTAATCCAGAGAACTAGAGCCCAACAAAATCTGCTTCAAGAGAAATTTCT 540  
QY 541 TGGCCCAAGAGAGCCGACACAGACATCACTTACATCCAAAAGGTCTCTCATCTCA 600

Db 541 TGGCCCAAGAGAGCCGACACAGACATCACTTACATCCAAAAGGTCTCTCATCTCA 600  
QY 601 ACAGAGAGTGGCTCCACAGAGAGACCTGGGTAGGAAAAATCTATTAACATAAGAGA 660  
Db 601 ACAGAGAGTGGCTCCACAGAGAGACCTGGGTAGGAAAAATCTATTAACATAAGAGA 660  
QY 661 GCCAGCATTGAGACAAAAGCATGTAGCCAGCCAGGGCTGGGCCAGAGAGCCACTTC 720  
Db 661 GCCAGCATTGAGACAAAAGCATGTAGCCAGCCAGGGCTGGGCCAGAGAGCCACTTC 720  
QY 721 ACCTCAACAAAGAGCTGATCAACACCTGGCCCGCCAGGCTAAACCTGGAGCCAAAAGGA 780  
Db 721 ACCTCAACAAAGAGCTGATCAACACCTGGCCCGCCAGGCTAAACCTGGAGCCAAAAGGA 780  
QY 781 GCCATCTGCCCGACCTGATCTACATCCCAAGAGACCTGAAACATCAAGACAAGAAC 840  
Db 781 GCCATCTGCCCGACCTGATCTACATCCCAAGAGACCTGAAACATCAAGACAAGAAC 840  
QY 841 AACGCCAGTCCAGGAGCCAAAATCCAAAGAGGAGTCTTGACAGAGCTGGGATTTCTAAC 900  
Db 841 AACGCCAGTCCAGGAGCCAAAATCCAAAGAGGAGTCTTGACAGAGCTGGGATTTCTAAC 900  
QY 901 AAAACTTCAGGAACATCTCATACAGGATCAGCCCTAGAGTGGAGAGCACTTTCTAGTG 960  
Db 901 AAAACTTCAGGAACATCTCATACAGGATCAGCCCTAGAGTGGAGAGCACTTTCTAGTG 960  
QY 961 GGTGCGAGATTCGAGTCAAGATCAGATGTGGATATATCTTCAGACACAGATTCCTCAGC 1020  
Db 961 GGTGCGAGATTCGAGTCAAGATCAGATGTGGATATATCTTCAGACACAGATTCCTCAGC 1020  
QY 1021 CACGATGGTGGAAATGTGGGCCAGGAGTGAAGCTAGGCTTCAAGAAAAATCTGGTTA 1080  
Db 1021 CACGATGGTGGAAATGTGGGCCAGGAGTGAAGCTAGGCTTCAAGAAAAATCTGGTTA 1080  
QY 1081 TAAAGTATGTAGATTAAGTGGAGACGTGCGCACATGAGAAAACAGATGCTGGAAATCA 1140  
Db 1081 TAAAGTATGTAGATTAAGTGGAGACGTGCGCACATGAGAAAACAGATGCTGGAAATCA 1140  
QY 1141 CAGACACTACAGAGATACAGCCCTCAAGGCTCATCCAAACATGAGCTGGCAATAGAC 1200  
Db 1141 CAGACACTACAGAGATACAGCCCTCAAGGCTCATCCAAACATGAGCTGGCAATAGAC 1200  
QY 1201 ACAGTGTGTGTACTGTGGCGGAGACAATAGCTTCTTCTCAGCCAGGATCTCTGG 1260  
Db 1201 ACAGTGTGTGTACTGTGGCGGAGACAATAGCTTCTTCTCAGCCAGGATCTCTGG 1260  
QY 1261 GGAAGAGGCCAGCGGCTGTCAAGGCTTTTGGCGGTGACGGGACAGCGGCTGGGCT 1320  
Db 1261 GGAAGAGGCCAGCGGCTGTCAAGGCTTTTGGCGGTGACGGGACAGCGGCTGGGCT 1320  
QY 1321 GGAGCGGCGCTGGGCGGCTGTGGGTGAGGCACTTTGACCTGAGACCCAGAGAC 1380  
Db 1321 GGAGCGGCGCTGGGCGGCTGTGGGTGAGGCACTTTGACCTGAGACCCAGAGAC 1380  
QY 1381 ACCGCGCAACGGGTACCGAGCTAGTGCACACAGCCCGCTGCTGGCGCACTCTCT 1440  
Db 1381 ACCGCGCAACGGGTACCGAGCTAGTGCACACAGCCCGCTGCTGGCGCACTCTCT 1440  
QY 1441 GCACAAATCCGCTATGTGGCTCCAAACCGCGGACACATCTTCCGACCAAGCCACAA 1500  
Db 1441 GCACAAATCCGCTATGTGGCTCCAAACCGCGGACACATCTTCCGACCAAGCCACAA 1500  
QY 1501 CTTGGCGAGTGTAGAGGCTTACTTGGCTCCCTCACCAGACTCCGCTGTGGTACTCTA 1560  
Db 1501 CTTGGCGAGTGTAGAGGCTTACTTGGCTCCCTCACCAGACTCCGCTGTGGTACTCTA 1560  
QY 1561 CGCCAGGCGCTGTGTGTTACAAATCGCCGGGGGTACTCTTCTTGAAGGCGACAGAGG 1620  
Db 1561 CGCCAGGCGCTGTGTGTTACAAATCGCCGGGGGTACTCTTCTTGAAGGCGACAGAGG 1620  
QY 1621 GGTCAACCGCGACTCTCTCGGAGATGTACAGGCTGATTAAGAGAGTCTCTATGCGCG 1680  
Db 1621 GGTCAACCGCGACTCTCTCGGAGATGTACAGGCTGATTAAGAGAGTCTCTATGCGCG 1680

**THIS PAGE BLANK (USPTO)**

Db 1621 GCTCACCGCCGACTTCCTCCGGAGTATGTCACGCTGCATAGAGATGCTTATGCGCC 1680  
QY 1681 CTGCCTGGGCTTCAGTTCAAGTCAGGCTGCATCCGGCCATTCTCTGCAGACATCTCATTTG 1740  
Db 1681 CTGCCTGGGCTTCAGTTCAAGTCAGGCTGCATCCGGCCATTCTCTGCAGACATCTCATTTG 1740  
QY 1741 GCTGCTGCTCTTGGGGGAGCACTACAAAGCAGACAGAGAGGCTCAGTGTGGCCGAC 1800  
Db 1741 GCTGCTGCTCTTGGGGGAGCACTACAAAGCAGACAGAGAGGCTCAGTGTGGCCGAC 1800  
QY 1801 CTCTCTCTTACACGAGGCGGCTTGGCATGACCCGAGGCTGGTGGGCTGAGTTGA 1860  
Db 1801 CTCTCTCTTACACGAGGCGGCTTGGCATGACCCGAGGCTGGTGGGCTGAGTTGA 1860  
QY 1861 GCGGATACACAGAACCTTGGAGCTGACCTTCTGAAAAGCTTCTGGAACATCACGAGAT 1920  
Db 1861 GCGGATACACAGAACCTTGGAGCTGACCTTCTGAAAAGCTTCTGGAACATCACGAGAT 1920  
QY 1921 GGAAGTCTATGCTCTTGGGCAACATGCGATCGGCCACGCTGAGGGTAAAGCCGCTGCT 1980  
Db 1921 GGAAGTCTATGCTCTTGGGCAACATGCGATCGGCCACGCTGAGGGTAAAGCCGCTGCT 1980  
QY 1981 CACCTGACACCCGAGAACCTTTGATGATGACCTGACCTGACGACCCGACGCTCAGAGTCAC 2040  
Db 1981 CACCTGACACCCGAGAACCTTTGATGATGACCTGACCTGACGACCCGACGCTCAGAGTCAC 2040  
QY 2041 CATCTCACCCCACTGGCCACACAGGCTTGGGCCCTGCTCTGCTCAGAGCTCATCTCTTA 2100  
Db 2041 CATCTCACCCCACTGGCCACACAGGCTTGGGCCCTGCTCTGCTCAGAGCTCATCTCTTA 2100  
QY 2101 TGAACCTGCTGAAAGAGACAGACAGTGAAGCTGACAGCTGTAAAGTCCAAAGGCCA 2160  
Db 2101 TGAACCTGCTGAAAGAGACAGACAGTGAAGCTGACAGCTGTAAAGTCCAAAGGCCA 2160  
QY 2161 ACAGAGCTGAGAGCTGTGGCGCGGCCCCAGAGGACCCCGCTGGGGTCCCTGATAGT 2220  
Db 2161 ACAGAGCTGAGAGCTGTGGCGCGGCCCCAGAGGACCCCGCTGGGGTCCCTGATAGT 2220  
QY 2221 GCACTTCCACAGGCGGTGGCTTTGTGGCCAGACTCCAGATCCAGACAGGCTCACTCAA 2280  
Db 2221 GCACTTCCACAGGCGGTGGCTTTGTGGCCAGACTCCAGATCCAGACAGGCTCACTCAA 2280  
QY 2281 GAGCTGGGCGCCAGAGAGCTGGGGCGCCCATCATCTCATGACTCATCTCTGGCCCTGA 2340  
Db 2281 GAGCTGGGCGCCAGAGAGCTGGGGCGCCCATCATCTCATGACTCATCTCTGGCCCTGA 2340  
QY 2341 GGGCCCTTCCCGCGTGGCGCTGGAGAGTGTCTTGGCTTACTGCTGGGCGCATCAAGCA 2400  
Db 2341 GGGCCCTTCCCGCGTGGCGCTGGAGAGTGTCTTGGCTTACTGCTGGGCGCATCAAGCA 2400  
QY 2401 CTGCGGCTCTCTTGGCTCAACAGAGGGAACGATCTGCTTGGCGGGGAGAGTGCAGGCG 2460  
Db 2401 CTGCGGCTCTCTTGGCTCAACAGAGGGAACGATCTGCTTGGCGGGGAGAGTGCAGGCG 2460  
QY 2461 GAACTCTGCTTACAGCTGTGCTCTTGGGAGAGAGCTTACGAGGAGTGGGGTGCAGATGG 2520  
Db 2461 GAACTCTGCTTACAGCTGTGCTCTTGGGAGAGAGCTTACGAGGAGTGGGGTGCAGATGG 2520  
QY 2521 CATCATGAGAGCTTACCGGCGCAATGCTGACGCGCTGCGGCTCTCCCTCCGCGCTGCT 2580  
Db 2521 CATCATGAGAGCTTACCGGCGCAATGCTGACGCGCTGCGGCTCTCCCTCCGCGCTGCT 2580  
QY 2581 GAGGCTCATGAGACCTTGTGCTCCCTCACTGCTCTCCAAAGTGTCAAGCGCTATGTC 2640  
Db 2581 GAGGCTCATGAGACCTTGTGCTCCCTCACTGCTCTCCAAAGTGTCAAGCGCTATGTC 2640  
QY 2641 TGGTGAAGAAGAGAGAGACACATCCAACTAGACAGACAGAAAGCCCTCGGCATGATGGGCT 2700  
Db 2641 TGGTGAAGAAGAGAGAGACACATCCAACTAGACAGACAGAAAGCCCTCGGCATGATGGGCT 2700  
QY 2701 GGTGCGGCGGAGACAGACCTGCTCTCCGAGACTTCGCGCTGGGCTGCTCATGAGCT 2760  
Db 2701 GGTGCGGCGGAGACAGACCTGCTCTCCGAGACTTCGCGCTGGGCTGCTCATGAGCT 2760

QY 2761 CAACTCCTTCTGAGTTAAAGTGGGCGCAAGTCCAGAGAAGATGTGAGACCCATAGCAGA 2820  
Db 2761 CAACTCCTTCTGAGTTAAAGTGGGCGCAAGTCCAGAGAAGATGTGAGACCCATAGCAGA 2820  
QY 2821 GCCGATGCGCGCAGTGTGTGAAGCAGACACTGGCCAGGCCAGGGCCCACTGGGAC 2880  
Db 2821 GCCGATGCGCGCAGTGTGTGAAGCAGACACTGGCCAGGCCAGGGCCCACTGGGAC 2880  
QY 2881 GGATTCCCTCAAGAACCTGAGGAGCTTGAAGGCTTGAAGGGAAGAACTCCGAGACTG 2940  
Db 2881 GGATTCCCTCAAGAACCTGAGGAGCTTGAAGGCTTGAAGGGAAGAACTCCGAGACTG 2940  
QY 2941 GTGAGACACCCCGAGATGTGCTGCTGAGAGACTTGAAGCCCTCAGACCTCCGCA 3000  
Db 2941 GTGAGACACCCCGAGATGTGCTGCTGAGAGACTTGAAGCCCTCAGACCTCCGCA 3000  
QY 3001 TGTCAACTTCTTATTAACCACTGAGAGATGACAGGGGAGAGGCTGAGCCAAAATGAGCT 3060  
Db 3001 TGTCAACTTCTTATTAACCACTGAGAGATGACAGGGGAGAGGCTGAGCCAAAATGAGCT 3060  
QY 3061 GAGGCCCATGAGACAGAGGCTGGGGTCCGTGCGGCTTCCCGAGGGTTCCACCCCG 3120  
Db 3061 GAGGCCCATGAGACAGAGGCTGGGGTCCGTGCGGCTTCCCGAGGGTTCCACCCCG 3120  
QY 3121 ACGCTCAGCCAGGGTGGCCACACAGATGCCCTTACTCTACACCATAGTCAAGAACCC 3180  
Db 3121 ACGCTCAGCCAGGGTGGCCACACAGATGCCCTTACTCTACACCATAGTCAAGAACCC 3180  
QY 3181 CTTGATGTGCGCGCTGCTGCGACCCGACAGCATGCTCAAGAGCCCTGACCTGTGCACAT 3240  
Db 3181 CTTGATGTGCGCGCTGCTGCGACCCGACAGCATGCTCAAGAGCCCTGACCTGTGCACAT 3240  
QY 3241 CGTGGCGTGGGCGGTGAGACCCCATGCTGAGACATCTGGTTCATGCTGCGGGCGGACTCG 3300  
Db 3241 CGTGGCGTGGGCGGTGAGACCCCATGCTGAGACATCTGGTTCATGCTGCGGGCGGACTCG 3300  
QY 3301 CAACTGCGGCGGCGGTGAGCTGCGGCTGGTGGAGACCTTGCAGCGGCTTCTTCTGAC 3360  
Db 3301 CAACTGCGGCGGCGGTGAGCTGCGGCTGGTGGAGACCTTGCAGCGGCTTCTTCTGAC 3360  
QY 3361 CCTAGCGGCGGTGAGCTGCGGCGGAGCGCGCAGAGCTGTGCTGAGCGCATCG 3420  
Db 3361 CCTAGCGGCGGTGAGCTGCGGCGGAGCGCGCAGAGCTGTGCTGAGCGCATCG 3420  
QY 3421 CCTGCTCTCACTCTCTCCCGCGGAGCCGAGCGGCGGAGAGCGGGGCTGCGGGGT 3480  
Db 3421 CCTGCTCTCACTCTCTCCCGCGGAGCCGAGCGGCGGAGAGCGGGGCTGCGGGGT 3480  
QY 3481 AGAGGGGCGTGGCGGCGGAGCACTAAAGCTGTGTTGCCATCTGCGCGGCTCG 3540  
Db 3481 AGAGGGGCGTGGCGGCGGAGCACTAAAGCTGTGTTGCCATCTGCGCGGCTCG 3540  
QY 3541 TCATGAATGCTTCCGGGCGGGGCGGAGAGGAGCGGGCTGTGCTTAAGTCGGG 3600  
Db 3541 TCATGAATGCTTCCGGGCGGGGCGGAGAGGAGCGGGCTGTGCTTAAGTCGGG 3600  
QY 3601 GTGGCAAGGGGCGGGGCGGGGCGGAAAGCTGAGACCTTGCACAGCGGGAGGGGAGCG 3660  
Db 3601 GTGGCAAGGGGCGGGGCGGGGCGGAAAGCTGAGACCTTGCACAGCGGGAGGGGAGCG 3660  
QY 3661 GCACACACACGCTGACCGAGACGCTGAGACTGACAGCCACGCTGCTTGGCTGCTG 3720  
Db 3661 GCACACACACGCTGACCGAGACGCTGAGACTGACAGCCACGCTTGGCTGCTG 3720  
QY 3721 CTGCTGCGGCGACCGCGCAGAGAGGAGGAGCTGCGCTCTCCAGAGTGGTGGT 3780  
Db 3721 CTGCTGCGGCGACCGCGCAGAGAGGAGGAGCTGCGCTCTCCAGAGTGGTGGT 3780  
QY 3781 GTTGAATAAAGTATTTAATTA 3804  
Db 3781 GTTGAATAAAGTATTTAATTA 3804

**THIS PAGE BLANK (USPTO)**



score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

(without alignments)  
17297.980 Million cell updates/sec

Sequence: 1 cttctgtagaagtgtcta.....taataaagtatttaatta 3804

Scoring table: IDENTITY\_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

GenEmb1: \*

- 1: gb\_ba:\*
- 2: gb\_hhg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_jm:\*
- 20: em\_om:\*
- 21: em\_ov:\*
- 22: em\_or:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_hhg\_hum:\*
- 31: em\_hhg\_in:\*
- 32: em\_hhg\_other:\*
- 33: em\_hhg\_mus:\*
- 34: em\_hhg\_pln:\*
- 35: em\_hhg\_rod:\*
- 36: em\_hhg\_nam:\*
- 37: em\_hhg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hhg\_hum:\*
- 40: em\_hhg\_mus:\*
- 41: em\_hhg\_other:\*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query	Match	length	DB	ID	Description
1	3804	100.0	3804	6	AX113509	AX113509 Sequence	
2	3804	100.0	3804	9	HSU40002	U40002 Human hormo	
3	2650.4	69.7	3255	9	HUMHSLA	L11706 Human hormo	
4	1896.4	49.9	3704	10	RNU6U0001	U40001 Rattus norv	
5	1841.8	48.4	2792	10	AF1774401	AF1774401 Spermoph	
6	1796.6	47.2	3125	4	SSHSITPAS	AJ000482 Sus scrof	
7	1636.4	43.0	3226	10	RATPIP	XS1415 Rattus sp.	
8	1620	42.6	2678	10	BC021642	BC021642 Mus muscu	
9	1619.2	42.6	3172	10	BM008188	U08188 Mus muscu	
10	1555.2	30.4	168586	9	AC011497	AC011497 Homo sapi	
11	1552.8	14.5	652	9	AY063230	AY063230 Canis fam	
12	538.2	14.1	656	4	AF141958	AF141958 Sus scrof	
13	531.6	14.0	140808	8	AC021858	AC021858 Homo sapi	
14	497.2	13.1	177575	2	AC021857	AC021857 Homo sapi	
15	436	11.5	2222	4	SSHS45UPR	AJ000483 Sus scrof	
16	405.6	10.7	498	4	SSAU61692	AJ224692 Sus scrof	
17	392.8	10.3	17835	10	AF179427	AF179427 Mus muscu	
18	392.8	10.3	190860	2	AC073803	AC073803 Mus muscu	
19	390.2	10.3	215234	2	AC073745	AC073745 Mus muscu	
20	388.6	10.2	215063	2	AC130942	AC130942 Rattus no	
21	388.6	10.2	155817	2	AC121639	AC121639 Rattus no	
22	282.8	7.4	142394	2	AC109038	AC109038 Rattus no	
23	272	7.2	4911	4	SSC6076	AJ006076 Sus scrof	
24	269.6	7.1	131063	2	AC130942	AC130942 Rattus no	
25	258	6.8	258	11	GI15903	GI15903 human SYS C	
26	256.6	6.7	311	4	BTA237675	AJ237675 Bos tauru	
27	256.4	6.7	155817	2	AC121639	AC121639 Rattus no	
28	253	6.7	429	4	BUT78042	U78042 Bos taurus	
29	245.4	6.5	1513	4	SSC6075	AJ006075 Sus scrof	
30	225.4	5.9	34918	3	AC004321	AC004321 Drosophi	
31	225.4	5.9	118533	2	AC020205	AC020205 Drosophi	
32	225.4	5.9	180814	3	AC099023	AC099023 Drosophi	
33	225.4	5.9	183592	3	AC099012	AC099012 Drosophi	
34	225.4	5.9	256764	3	AE003792	AE003792 Drosophi	
35	186.6	4.9	225	4	S80110	S80110 hormone-sen	
36	171	4.5	190860	4	AC073803	AC073803 Mus muscu	
37	149.8	3.9	2328	3	AY051867	AY051867 Drosophi	
38	112.8	3.0	142394	2	AC109038	AC109038 Rattus no	
39	102	2.7	106209	1	D90903	D90903 Synecocyst	
40	99.8	2.6	2359	1	PFPAIRAA	M17877 P. falciparu	
41	96.6	2.5	3187	6	A00663	A00663 P. falciparu	
42	96.4	2.5	137721	2	AC113928	AC113928 Rattus no	
43	94.6	2.5	1471	8	ZMIPROPR	Y17332 Zea mays nr	
44	94.6	2.5	1792	8	ZMAI0830	AJ130830 Zea mays	
45	87.4	2.3	4186	3	AY061106	AY061106 Drosophi	

ALIGNMENTS

RESULT 1

LOCUS AX113509

DEFINITION Sequence 1 from Patent WO0126664.

ACCESSION AX113509

VERSION AX113509.1

KEYWORDS GI:13939727

SOURCE human.

ORGANISM Homo sapiens

EuKaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 3804)

AUTHORS Mitchell,G.A. and Wang,S.P.

TITLE Hormone-sensitive lipase activity mediated male infertility

JOURNAL Patent: WO 0126664-A 1 19-APR-2001;

Dp	541	TCGCCACAGAGAGCGCCGACCACACAGCAATATCCTTACATCCAAAGGGTGTGCTCACC	600
Qy	601	ACAGAGAGCTCCTCCAGCAGAGGACCTGGGCTAGGAAAAAGATCTATACCTCAACAGGA	660
Dp	601	ACAGAAAGCTCCCTCCAGCAGGAGCCTGGGCTAGGAAAAAGATCTATACCTCAACAGGA	660
Qy	661	GCAGCAGTGTGACAAAGACATGTAGCCACGACGAGGCGCTGGGGCAGAGAGCCACTCC	720
Dp	661	GCAGCAGTGTGACAAAGACATGTAGCCACGACGAGGCGCTGGGGCAGAGAGCCACTCC	720
Qy	721	AGCTCAACAAGAACTGAAATCAACACCTGCGGCCAGGCTAAACCTGAGCAAGCAAGGA	780
Dp	721	AGCTCAACAAGAACTGAAATCAACACCTGCGGCCAGGCTAAACCTGAGCAAGCAAGGA	780
Qy	781	GCATATCGCCCGGAGCTGAAATCTATCAATCCCAAGACACCTGACAGCTGACAGCAAC	840
Dp	781	GCATATCGCCCGGAGCTGAAATCTATCAATCCCAAGACACCTGACAGCTGAGCAAGCAAC	840
Qy	841	AACGCCAGTCCAGGAGGAGCCAAATCCAAAGCAGGAGATCTTTGACAGAGCTGGGATTTTAC	900
Dp	841	AACGCCAGTCCAGGAGGAGCCAAATCCAAAGCAGGAGATCTTTGACAGAGCTGGGATTTTAC	900
Qy	901	AAACCTTCAGAGACTATCCATACAGCGATCAGCCCTAGAGTGGAAAGGCACTTTCTGAGTg	960
Dp	901	AAACCTTCAGAGAACTATCCATACAGCGATCAGCCCTAGAGTGGAAAGGCACTTTCTGAGTg	960
Qy	961	GGTCGAGATTTGAGTCTGAGTCTGAGATTCAGATGTGGGATTCATCTTCAACACACATCTCCAG	1020
Dp	961	GGTCGAGATTTGAGTCTGAGTCTGAGATTCAGATGTGGGATTCATCTTCAACACACATCTCCAG	1020
Qy	1021	CACGATGGGTGGATGTGTGGCCAGGAGTGAAGCTGAGGCTTCAAAGAAATCTGTTA	1080
Dp	1021	CACGATGGGTGGATGTGTGGCCAGGAGTGAAGCTGAGGCTTCAAAGAAATCTGTTA	1080
Qy	1081	TAAAGTGAATGTAGATACAGTACAGTGGGAGAGTGGCCACATGAGAAAAACAGTCTGGGATCA	1140
Dp	1081	TAAAGTGAATGTAGATACAGTACAGTGGGAGAGTGGCCACATGAGAAAAACAGTCTGGGATCA	1140
Qy	1141	CAGACACTACCAAGATACAGGCTCAAGGCTTCATCACACATGAGGACCTGGCAGCAATGAC	1200
Dp	1141	CAGACACTACCAAGATACAGGCTTCATCACACATGAGGACCTGGCAGCAATGAC	1200
Qy	1201	ACAGTGGTGGTGAATCTGGGGGAGGAGCAACATATGAGCTTTCTTGCAGCCAGGATCTGG	1260
Dp	1201	ACAGTGGTGGTGAATCTGGGGGAGGAGCAACATATGAGCTTTCTTGCAGCCAGGATCTGG	1260
Qy	1261	GGAAACGGGCCAGCGGCTGTAGGCGTTTTTGGCGGATGTACGGAGCAGAGCGCTGGGGCT	1320
Dp	1261	GGAAACGGGCCAGCGGCTGTAGGCGTTTTTGGCGGATGTACGGAGCAGAGCGCTGGGGCT	1320
Qy	1321	GGAGCGGGCGCGGGGCGGCGCTGTGGGTGTGGGCTGTGACCTTGACCTGAGCCAGAGAC	1380
Dp	1321	GGAGCGGGCGCGGGGCGGCGCTGTGGGTGTGGGCTGTGACCTTGACCTGAGCCAGAGAC	1380
Qy	1381	ACCGGCAACGGGTACCGCAGACCTAGTGCACACAGCGCCCTGTGGCTGGGCGCACTTCT	1440
Dp	1381	ACCGGCAACGGGTACCGCAGACCTAGTGCACACAGCGCCCTGTGGCTGGGCGCACTTCT	1440
Qy	1441	GCACAAATTCGGGCTATGTGGCTTCACACGCGCGCAGCATCTTTCTTCCGACACGCAACA	1500
Dp	1441	GCACAAATTCGGGCTATGTGGCTTCACACGCGCGCAGCATCTTTCTTCCGACACGCAACA	1500
Qy	1501	CCTGGCGAGTGTGAGGCTACCTGGCTGGCTGCACACAGCTCGCGGTGGTGGTCTACTA	1560
Dp	1501	CCTGGCGAGTGTGAGGCTACCTGGCTGGCTGCACACAGCTCGCGGTGGTGGTCTACTA	1560
Qy	1561	CGCCCAAGCGCTGTGTTACCAATCGCGCGGGGTACTCTTTTGAAGGCGACGAGGG	1620
Dp	1561	CGCCCAAGCGCTGTGTTACCAATCGCGCGGGGTACTCTTTTGAAGGCGACGAGGG	1620
Qy	1621	GCTACCGCGCACTTCCTCGGGAGTATGACGCTGATAAAGGATGCTTCTATGGCGG	1680

Db 1621 GCTACACGCCGACTCCCTCCGGGAGTATGTCACGCTGCATTAAGGATGCTTCTATGCGC 1680  
 QY 1681 CTGCGTGGGCTTCAGTTACAGGCTGCATCGCGCCATTCTCTGAGACCAATTCATTTGG 1740  
 Db 1681 CTGCGTGGGCTTCAGTTACAGGCTGCATCGCGCCATTCTCTGAGACCAATTCATTTGG 1740  
 QY 1741 GCTGTGTCCTTCGGGGAGCACTACAAACGCAAGAGAGAGGCTCTAGTGTGGCGCGCAG 1800  
 Db 1741 GCTGTGTCCTTCGGGGAGCACTACAAACGCAAGAGAGAGGCTCTAGTGTGGCGCGCAG 1800  
 QY 1801 CTCTCTCTTACACAGCGCGCGCTTTGGCATGACCCCGAGCTGCGTGGGCTGAGTTGA 1860  
 Db 1801 CTCTCTCTTACACAGCGCGCGCTTTGGCATGACCCCGAGCTGCGTGGGCTGAGTTGA 1860  
 QY 1861 GCGGATCACACAGACCTGGAGCTGCACCTCTGAAAAGCTTGGAAACATCACCGGAT 1920  
 Db 1861 GCGGATCACACAGACCTGGAGCTGCACCTCTGAAAAGCTTGGAAACATCACCGGAT 1920  
 QY 1921 GGAAGTGTATGCTCTCTGGCCAACTGGCATCGCCACCGTGAAGGGTAAAGCCGCTGCT 1980  
 Db 1921 GGAAGTGTATGCTCTCTGGCCAACTGGCATCGCCACCGTGAAGGGTAAAGCCGCTGCT 1980  
 QY 1981 CAGCGTCCACCCGAAACCTTTAGATGCCACTGACTGCGAGCCCAAGCTCACGGTAC 2040  
 Db 1981 CAGCGTCCACCCGAAACCTTTAGATGCCACTGACTGCGAGCCCAAGCTCACGGTAC 2040  
 QY 2041 CATCTACACCCCACTGGCCACACAGGCGCTGGGCGCGTCTGTCAGGCTCATCTCTTA 2100  
 Db 2041 CATCTACACCCCACTGGCCACACAGGCGCTGGGCGCGTCTGTCAGGCTCATCTCTTA 2100  
 QY 2101 TGAACCTGCGTGAAGGACAGAGACAGTGAAGAGCTCAGCAGCGCTGTAAGTCAACAGGCA 2160  
 Db 2101 TGAACCTGCGTGAAGGACAGAGACAGTGAAGAGCTCAGCAGCGCTGTAAGTCAACAGGCA 2160  
 QY 2161 AGGAGAGCTTGAGAGCTTGCGCGCGCCCGCAGAGGACCCCGCTCGCGTCCCTGATAGT 2220  
 Db 2161 AGGAGAGCTTGAGAGCTTGCGCGCGCCCGCAGAGGACCCCGCTCGCGTCCCTGATAGT 2220  
 QY 2221 GCACCTTCACGAGGCGTGGGCTTGTGGCCAGACCTCCAGATCCAGAGGCTCAACCTCAA 2280  
 Db 2221 GCACCTTCACGAGGCGTGGGCTTGTGGCCAGACCTCCAGATCCAGAGGCTCAACCTCAA 2280  
 QY 2281 GAGCTGGGCCACAGAGAGCTGGGCGCCCGCATCATCTCATGACTACCTCGCGCCCTGA 2340  
 Db 2281 GAGCTGGGCCACAGAGAGCTGGGCGCCCGCATCATCTCATGACTACCTCGCGCCCTGA 2340  
 QY 2341 GCGCCCTTCCCGCGTGGCGTGGAGAGTCTTCTGCTACTGCTGGGCTCAAGCA 2400  
 Db 2341 GCGCCCTTCCCGCGTGGCGTGGAGAGTCTTCTGCTACTGCTGGGCTCAAGCA 2400  
 QY 2401 CTGCGGCTCTCTTGGCTCAACAGGGGAAAGATCTGCTGGGGGAGAGTGCAGGCGG 2460  
 Db 2401 CTGCGGCTCTCTTGGCTCAACAGGGGAAAGATCTGCTGGGGGAGAGTGCAGGCGG 2460  
 QY 2461 GAACCTCTGCTTACCCGTGAGCTTCTGGGAGCAGAGCTTACGGGGTGGGGTGGCAGATGG 2520  
 Db 2461 GAACCTCTGCTTACCCGTGAGCTTCTGGGAGCAGAGCTTACGGGGTGGGGTGGCAGATGG 2520  
 QY 2521 CATCATGAGAGCTTACCCGCGCCAAATGCTGAGCTGCGCGCTCTCCCGCGCTGCT 2580  
 Db 2521 CATCATGAGAGCTTACCCGCGCCAAATGCTGAGCTGCGCGCTCTCCCGCGCTGCT 2580  
 QY 2581 GAGCGCTCATGAGACCCCTTGTGCGCCCTCAGTGTGCTCTCAATGTGTGAGCGCTATGC 2640  
 Db 2581 GAGCGCTCATGAGACCCCTTGTGCGCCCTCAGTGTGCTCTCAATGTGTGAGCGCTATGC 2640  
 QY 2641 TGGTGAAGAGAGAGAGACACTTCAACTCAGACAGAAAGCCCTGCGCATGTGAGGGCT 2700  
 Db 2641 TGGTGAAGAGAGAGAGACACTTCAACTCAGACAGAAAGCCCTGCGCATGTGAGGGCT 2700  
 QY 2701 GGTGGGGGGGAGACAGCCCTGCTCTCGAGACTTCCGCTGGGTGCTCTCATGGCT 2760  
 Db 2701 GGTGGGGGGGAGACAGCCCTGCTCTCGAGACTTCCGCTGGGTGCTCTCATGGCT 2760

QY 2761 CAACCTCTCTGGAGTTAAAGTGGGCGCAAGTCCCGAAGATGTGAGAGCCCATGACAGA 2820  
 Db 2761 CAACCTCTCTGGAGTTAAAGTGGGCGCAAGTCCCGAAGATGTGAGAGCCCATGACAGA 2820  
 QY 2821 GCCATGCGCGCGAGTGTGTGAAGCAGACACTGGCCAGCCCGCAGGCGCCCATGGGAC 2880  
 Db 2821 GCCATGCGCGCGAGTGTGTGAAGCAGACACTGGCCAGCCCGCAGGCGCCCATGGGAC 2880  
 QY 2881 GGATTCCTTCAAGAACTGACCTGAGGAGCTTGAAGCTTGAAGGAAACTCCGAGAGCTC 2940  
 Db 2881 GGATTCCTTCAAGAACTGACCTGAGGAGCTTGAAGCTTGAAGGAAACTCCGAGAGCTC 2940  
 QY 2941 GTGAGACACCCCGAGATGTGCTGATGAGTACACTTTGAGCCCTCCACACCCCTCGA 3000  
 Db 2941 GTGAGACACCCCGAGATGTGCTGATGAGTACACTTTGAGCCCTCCACACCCCTCGA 3000  
 QY 3001 TGTCAACTTCTTATTAACCACTGAGAGATGACAGGGGAAAGAGCTGAGCCAAATGAGCT 3060  
 Db 3001 TGTCAACTTCTTATTAACCACTGAGAGATGACAGGGGAAAGAGCTGAGCCAAATGAGCT 3060  
 QY 3061 GAGCCCATGAGACAGAGGCTTGGGCTCGTGGCGCTTCCCGAGGTTTCCACCCCG 3120  
 Db 3061 GAGCCCATGAGACAGAGGCTTGGGCTCGTGGCGCTTCCCGAGGTTTCCACCCCG 3120  
 QY 3121 ACGCTCCAGCGAGGTTGCCACACAGATGCGCCCTGCTACCTCCTACCCATGATCAAGAACCC 3180  
 Db 3121 ACGCTCCAGCGAGGTTGCCACACAGATGCGCCCTGCTACCTCCTACCCATGATCAAGAACCC 3180  
 QY 3181 CTTCATGTGCGCGCTGTGCGCACCCGACAGCATGCTCAAGAGCCCTGCCACCTGTGACAT 3240  
 Db 3181 CTTCATGTGCGCGCTGTGCGCACCCGACAGCATGCTCAAGAGCCCTGCCACCTGTGACAT 3240  
 QY 3241 CGTGGCGTGGCGCTGGAACCCCATGTGAGACACTCGCTCATGCTCGCGCGCGCATGCG 3300  
 Db 3241 CGTGGCGTGGCGCTGGAACCCCATGTGAGACACTCGCTCATGCTCGCGCGCGCATGCG 3300  
 QY 3301 CAACCTGGGCGAGCGGAGTGGAGCTGGCGGCTGGTGAAGAACTGCCCGCAGGCTTCTCAG 3360  
 Db 3301 CAACCTGGGCGAGCGGAGTGGAGCTGGCGGCTGGTGAAGAACTGCCCGCAGGCTTCTCAG 3360  
 QY 3361 CCTAGCGCGCTGTGCGCGAGAGCGCGCAGAGCGCGCAAGAGCTGTGCTGAGAGCGCATCG 3420  
 Db 3361 CCTAGCGCGCTGTGCGCGAGAGCGCGCAGAGCGCGCGCAAGAGCTGTGCTGAGAGCGCATCG 3420  
 QY 3421 CTTGCTCTCACTCTCCCGCGAGACCGGCGCAGCGGGAGAGAGCGGGGCTGCGGGGCT 3480  
 Db 3421 CTTGCTCTCACTCTCCCGCGAGACCGGCGCAGCGGGAGAGAGCGGGGCTGCGGGGCT 3480  
 QY 3481 AGAGGGGGGCGTGGGGGGGAGCACTAAAGCTGTGTTTCCATCTGCGCGCGGCTCGG 3540  
 Db 3481 AGAGGGGGGCGTGGGGGGGAGCACTAAAGCTGTGTTTCCATCTGCGCGGCTCGG 3540  
 QY 3541 TCATGAATGCTTCCGGGCGCGGCGGAGAGGCGGCTGTGCTTAAAGTGGGG 3600  
 Db 3541 TCATGAATGCTTCCGGGCGCGGCGGAGAGGCGGCTGTGCTTAAAGTGGGG 3600  
 QY 3601 GTGCAAGAGGGGGCGGGGGGCGGAAAGCTGAGACCTTGGCCACAGGGGAGGGGAGCGC 3660  
 Db 3601 GTGCAAGAGGGGGCGGGGGGCGGAAAGCTGAGACCTTGGCCACAGGGGAGGGGAGCGC 3660  
 QY 3661 GCACACACACCGGTACCGAGACAGGCTGAGCTGACAGCCACCGCTGCTTGTGCTG 3720  
 Db 3661 GCACACACACCGGTACCGAGACAGGCTGAGCTGACAGCCACCGCTGCTTGTGCTG 3720  
 QY 3721 CTGCTGGCGGACCGCGCAGAGGAGCGGGAGCTGCGCTTCCCTGAGAGTTCGGTTG 3780  
 Db 3721 CTGCTGGCGGACCGCGCAGAGGAGCGGGAGCTGCGCTTCCCTGAGAGTTCGGTTG 3780  
 QY 3781 GTTGAATAAAGTATTTAATTA 3804  
 Db 3781 GTTGAATAAAGTATTTAATTA 3804

```

RESULT 2
LOCUS HSU40002
DEFINITION Human hormone-sensitive lipase testicular isoform mRNA, complete
ACCESSION U40002
VERSION 040002.1
KEYWORDS GI:1488676
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 3804)
AUTHORS Holst,L.S., Langin,D., Mulder,H., Laurell,H., Grober,J., Bergh,A.,
Mohenwieser,H.W., Edgren,G. and Holm,C.
Molecular cloning, genomic organization, and expression of a
testicular isoform of hormone-sensitive lipase
JOURNAL Genomics 35 (3), 441-447 (1996)
MEDLINE 97001144
PUBMED 8812477
REFERENCE 2 (bases 1 to 3804)
AUTHORS Stenson Holst,L., Langin,D., Laurell,H., Grober,J., Edgren,G. and
Holm,C.
Direct Submission
Submitted (02-NOV-1995) Lena Stenson Holst, Cell and Molecular
Biology, Lund University, P.O. Box 94, Lund, S-22100, Sweden
Location/Qualifiers
1. 3804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.1-13.2"
/tissue_type="testis"
278..3508
/codon_start=1
/product="hormone-sensitive lipase testicular isoform"
/protein_id="AAC5066.1"
/db_xref="GI:1488677"
/translating="MEPSKSVSRSDMPPEHPORITPLEPEKPTPIAPOESKTLG
SNTQOKPASNOPLTOETPAOHDSQEKERAOOKSJOEPLAPOKPAOQSPYIO
RVLLQOEAASQDGLGKESITQOEPALORHNAOPGGEPPPAOEBSTPAAO
AKPGKREPSAPTESTSOETPEOSDKOTTPYQAGSKOSJLLEGLFKLELSORS
ALEBKALSEYWADESBSVDSDSDPATMGAVAGVVLGKSGYVMGYSG
TSPEKTSARHRYDYDASRLINMDLRTQSLVLAEDNIAFFSGSGETAORL
SGVAGVREOALGLEPAGRLGVLAHLFDLDETPANGVRSIVTARCLHLHKS
YVANSRSIFERTSHNLAEVLAALTOULALVYAGRLVTRNPGVLFEEGEGLT
ADFLREVTLHKGCEYGRGFOETPAIRPLIOTISGIVSFGHYKRNELGSVAAS
SLTSGRAIDPELGAPEERITONLDVHPKATNTEMYVLSLANMASIVRSR
LSTLPPRAPEPLTADTLTYTISPPLAHTEGVLVRLISYDLREGDSELSLIK
SNGORSLELMPPOAPRSRLIVHFGGGAOVSRSHEPYLSMGOELGAPLISID
YSLAPEAPFPALECEFFAYCMAIKHCLALSTGERICLADSGAGNCFVALRAAA
YGVAVPGIIMAAVPMALOPASPSRLISLPLSVLSKCSAAGATTEHNS
DOKALGMGIVRDTALLRDLRGLGASMSLSFELSGRSKQVSAEPMRVSVE
AALAOPOGLTGLDLSLNLRLDLSLNGNSNSPTPELSAFTLSESTPBDVNFLLP
PEDAGEAEAKNELSPMDRGIVRAAPREGHPRSSOGATOMPLISPIVKNFEMP
ILADSDMSLPPHYIYACALDPLDSDVMILARLNLGQVTLRVVDEPHGLTLA
ALCRETQOAEELCVERIRLVITPPAGAGPSETGAAGVGGCGGRH"
BASE COUNT 828 a 1225 c 1091 g 660 t
ORIGIN
Query Match 100.0%; Score 3804; DB 9; Length 3804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

121 CTTTAAAGAGAGACAGTGTAGATGAGTGAAGTGAATCCCAAGAGAAAGTCCATCA 180
121 CTTTAAAGAGAGACAGTGTAGATGAGTGAAGTGAATCCCAAGAGAAAGTCCATCA 180
181 TGAGGAATCAATAGAGATCTGTGAAGAGAGAGAGCGTGGGTGGAGCCCAAGAGATAGA 240
181 TGAGGAATCAATAGAGATCTGTGAAGAGAGAGAGCGTGGGTGGAGCCCAAGAGATAGA 240
241 ACCTGGAATCAATATCTCCGTGAGGGAATACATATGAGCCAGGTTCTTAAAGTCACT 300
241 ACCTGGAATCAATATCTCCGTGAGGGAATACATATGAGCCAGGTTCTTAAAGTCACT 300
301 GTCTAGCTAGACTGCAACTGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
301 GTCTAGCTAGACTGCAACTGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
361 GCCAGAAAAGACACCCATAGCCAGCAGAGATGCAAGACTGTGACAGGATTCACCA 420
361 GCCAGAAAAGACACCCATAGCCAGCAGAGATGCAAGACTGTGACAGGATTCACCA 420
421 ACAGAAAGCTGCTTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
421 ACAGAAAGCTGCTTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
481 TGCTGAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
481 TGCTGAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
541 TGCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 TGCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
601 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
601 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
661 GCCAGCATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
661 GCCAGCATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
721 AGCTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 AGCTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
781 GCCATCTGCCCCGAGTGAATCTACATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 840
781 GCCATCTGCCCCGAGTGAATCTACATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 840
841 AACGCCAGTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
841 AACGCCAGTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
901 AAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
901 AAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
961 GGTGCAATCTGAGTGAATCAGATGAGATGAGATCCTTCAAGAGAGAGAGAGAGAGAG 1020
961 GGTGCAATCTGAGTGAATCAGATGAGATGAGATCCTTCAAGAGAGAGAGAGAGAGAG 1020
1021 CAGATGGGTGAGTGGTGGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1021 CAGATGGGTGAGTGGTGGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1081 TAAAGTATGTCAAGATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
1081 TAAAGTATGTCAAGATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
1141 CAGACACTACAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1141 CAGACACTACAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1201 ACACTGCTGGTGAATCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1201 ACACTGCTGGTGAATCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

```

D	1201	ACAGTCGGCTGGAGCTCTGGCGGAGGACAACTAGCCCTTCCTTCAGGCCAGGGCTCTGG	1260
Q	1261	GGAAAGCGGCCACAGGGCTCTGACGGCGTTTTCGGCGGTGTACGGAGACAGGCGCTGGGGCT	1320
D	1261	GGAAAGCGGCCACAGGGCTCTGACGGCGTTTTCGGCGGTGTACGGAGACAGGCGCTGGGGCT	1320
Q	1321	GGAGCGGGCCCTTGGGCCGCGCTCTGGGTGTGGCGACCTTTTGACTGGACCTGGACCCAGAGAC	1380
D	1321	GGAGCGGGCCCTTGGGCCGCGCTCTGGGTGTGGCGACCTTTTGACTGGACCTGGACCCAGAGAC	1380
Q	1381	ACCGGCGCAACGGGGTACCGGAGGCTAGTGCACACAGCCCGCTGCTGGCTGGCGACCTCCT	1440
D	1381	ACCGGCGCAACGGGGTACCGGAGGCTAGTGCACACAGCCCGCTGCTGGCTGGCGACCTCCT	1440
Q	1441	GCACAAATCCCGCTATGTGGCCGCCAACCGCGGACAGCTATTTTTCGCGACCGGACAA	1500
D	1441	GCACAAATCCCGCTATGTGGCCGCCAACCGCGGACAGCTATTTTTCGCGACCGGACAA	1500
Q	1501	CGTGGCCGAGCTGAGAGGCTTACCTGGGCTCCCTCACCCAGCTCCGCGCTCTGGTCTACTA	1560
D	1501	CGTGGCCGAGCTGAGAGGCTTACCTGGGCTCCCTCACCCAGCTCCGCGCTCTGGTCTACTA	1560
Q	1561	CGCCACGCGCTCTGTGTGTACCAATCGCGGGGGTACTTTCTTTTGAGGGCGACGAGG	1620
D	1561	CGCCACGCGCTCTGTGTGTACCAATCGCGGGGGTACTTTCTTTTGAGGGCGACGAGG	1620
Q	1621	GCTCACCGCGGACTTCCTCCGGGAGATGTCAACGCTGCATTAAGGGATGCTCTCATATGGCC	1680
D	1621	GCTCACCGCGGACTTCCTCCGGGAGATGTCAACGCTGCATTAAGGGATGCTCTCATATGGCC	1680
Q	1681	CTGCGTGGGCTTCCAGTTTACGCTCTGCCATCCGGGCTTCCTGCAACCATCTCCATTGG	1740
D	1681	CTGCGTGGGCTTCCAGTTTACGCTCTGCCATCCGGGCTTCCTGCAACCATCTCCATTGG	1740
Q	1741	GCTGTGTCTTTCGGGGAGACTTACAACGCAACGAGACAGGCGCTCAGTGTGGCGCCAG	1800
D	1741	GCTGTGTGTCTTTCGGGGAGACTTACAACGCAACGAGACAGGCGCTCAGTGTGGCGCCAG	1800
Q	1801	CTCTCTTTCACACAGGGGCGGCTTTGGCATGCATGACCCCGAGCGCGGGGCTCAGTTTGA	1860
D	1801	CTCTCTTTCACACAGGGGCGGCTTTGGCATGCATGACCCCGAGCGCGGGGCTCAGTTTGA	1860
Q	1861	GCGGATTCACACAGAACTTGGAGCTGTGCACCTTGTGAAAGGCTTGTGAAACATCACAGAT	1920
D	1861	GCGGATTCACACAGAACTTGGAGCTGTGCACCTTGTGAAAGGCTTGTGAAACATCACAGAT	1920
Q	1921	GGAAGTGTATCTGTCTGGCCAACTGGCATCGGCCACGTTGAGGGTAAAGCGGCTGCT	1980
D	1921	GGAAGTGTATCTGTCTGGCCAACTGGCATCGGCCACGTTGAGGGTAAAGCGGCTGCT	1980
Q	1981	CAGCGTGCACCCGAAAGCCTTTGAGATGTCACCTGACTGGCCGAGCCCAAGCTCAGGGTAC	2040
D	1981	CAGCGTGCACCCGAAAGCCTTTGAGATGTCACCTGACTGGCCGAGCCCAAGCTCAGGGTAC	2040
Q	2041	CATCTCACCCCTACAGGAGGCGGCTGGGCGCGTCTGCTGTCAGGCTCATCTCTCTA	2100
D	2041	CATCTCACCCCTACAGGAGGCGGCTGGGCGCGTCTGCTGTCAGGCTCATCTCTCTCTA	2100
Q	2101	TGACCTTGGTGAAGACAGAGAGCTCAGAGAGCTGATTAAGTCAACGGGCA	2160
D	2101	TGACCTTGGTGAAGACAGAGAGCTCAGAGAGCTGATTAAGTCAACGGGCA	2160
Q	2161	ACGAGAGCTGAGAGCTGTGGCGCGCGCCGACAGAGGACCCCGCTCGCGTCTGTATAGT	2220
D	2161	ACGAGAGCTGAGAGCTGTGGCGCGCGCCGACAGAGGACCCCGCTCGCGTCTGTATAGT	2220
Q	2221	GCACCTTCCACGGGCTGTGGCTTTGTGGCCAGAGCTCCAGATCCCAAGAGCCCTTACTTCA	2280
D	2221	GCACCTTCCACGGGCTGTGGCTTTGTGGCCAGAGCTCCAGATCCCAAGAGCCCTTACTTCA	2280
Q	2281	GAGCTGGGCGCAGAGCTGGGCGCCCAATCTTCATCGACTACTCCCTGGCCCTCGA	2340
D	2281	GAGCTGGGCGCAGAGCTGGGCGCCCAATCTTCATCGACTACTCCCTGGCCCTCGA	2340
D	2341	GGCCCCCTTCCCGCGTGGAGAGTGTCTTTCGCCCTACTGCTGGGCGCATCAAGCA	2400
Q	2341	GGCCCCCTTCCCGCGTGGAGAGTGTCTTTCGCCCTACTGCTGGGCGCATCAAGCA	2400
D	2401	CTGGCGGCTCTTGGCTCAACAGGGGAAACGATCTGCTTCGCGGGACAGTGCAGCGCG	2460
Q	2401	CTGGCGGCTCTTGGCTCAACAGGGGAAACGATCTGCTTCGCGGGACAGTGCAGCGCG	2460
D	2401	CTGGCGGCTCTTGGCTCAACAGGGGAAACGATCTGCTTCGCGGGACAGTGCAGCGCG	2460
Q	2461	GAACCTCTGCTCAACGCTGGCTCTTGGGGGACAGAGCTAGGGGGTGTGGGGCGAGATGG	2520
D	2461	GAACCTCTGCTCAACGCTGGCTCTTGGGGGACAGAGCTAGGGGGTGTGGGGCGAGATGG	2520
Q	2521	CATCATGGAGGCTTACCGGGGCAACATATGTACAGCTTCCGCTCTCCCTCCGCTATGC	2580
D	2521	CATCATGGAGGCTTACCGGGGCAACATATGTACAGCTTCCGCTCTCCCTCCGCTATGC	2580
Q	2581	GAGCTTCATGAGCCCTTGTGCCCCCTCAGTGTGCTCCAAAGTGTCTCAGCGGCTATGC	2640
D	2581	GAGCTTCATGAGCCCTTGTGCCCCCTCAGTGTGCTCCAAAGTGTCTCAGCGGCTATGC	2640
Q	2641	TGTGTCAAAACGAGGAGACCACTCAACTCAGACACAAAGCCCTCGGCAATGATGGGCT	2700
D	2641	TGTGTCAAAACGAGGAGACCACTCAACTCAGACACAAAGCCCTCGGCAATGATGGGCT	2700
Q	2701	GGTGGCGGGGACACAGCCGCTCCGAGATTCCTGGGCTGGTGTCTCTCATGGCT	2760
D	2701	GGTGGCGGGGACACAGCCGCTCCGAGATTCCTGGGCTGGTGTCTCTCATGGCT	2760
Q	2761	CAACTCCCTTCTGAGTAAATGTGGGCGCCAAATGCCAAGATGTGCGAGGCCATAGAGA	2820
D	2761	CAACTCCCTTCTGAGTAAATGTGGGCGCCAAATGCCAAGATGTGCGAGGCCATAGAGA	2820
Q	2821	GCCGATGCGCCGACAGTGTCTGAGACGACATGCGCCAGCCGACAGGCGCCACTGGGAC	2880
D	2821	GCCGATGCGCCGACAGTGTCTGAGACGACATGCGCCAGCCGACAGGCGCCACTGGGAC	2880
Q	2881	GGAATCCCTCAAGAACCTGACCCGAGGAGGACTTGAAGCTGAGGGGAAACTCCGAGAGTC	2940
D	2881	GGAATCCCTCAAGAACCTGACCCGAGGAGGACTTGAAGCTGAGGGGAAACTCCGAGAGTC	2940
Q	2941	GTCGAGACCCCGAGATGTGCTGCTGACGTGACATTAAGCCCTTCACACCCCTCGA	3000
D	2941	GTCGAGACCCCGAGATGTGCTGCTGACGTGACATTAAGCCCTTCACACCCCTCGA	3000
Q	3001	TGTCACTCTTATATACCACTGAGATGCAAGGGGAAAGAGCTGAGGCCAAATATGAGCT	3060
D	3001	TGTCACTCTTATATACCACTGAGATGCAAGGGGAAAGAGCTGAGGCCAAATATGAGCT	3060
Q	3061	GAGCCCATGAGACAGAGGCGTGGGCGCGCTCCGAGGGGTTTCCACCCCG	3120
D	3061	GAGCCCATGAGACAGAGGCGTGGGCGCGCTCCGAGGGGTTTCCACCCCG	3120
Q	3121		

QY	3421	CCTGTCCTCACTCTCTCCCGCGGAGCGGCGGAGACGGGGGCTGCGGGGT	3480
Db	3421	CCTGTCCTCACTCTCTCCCGCGGAGCGGCGGAGACGGGGGCTGCGGGGT	3480
QY	3481	AGACGGGGGCTGCGGGGGCGACACTAAAGCCGTGTTCCATCTGCGCCGGCCCTCG	3540
Db	3481	AGACGGGGGCTGCGGGGGCGACACTAAAGCCGTGTTCCATCTGCGCCGGCCCTCG	3540
QY	3541	TCATGAATGCTTCCGGGGCGGGGAGGAGCGCGGCTGTGCTTACTTTAACTCGGG	3600
Db	3541	TCATGAATGCTTCCGGGGCGGGGAGGAGCGCGGCTGTGCTTACTTTAACTCGGG	3600
QY	3601	GTGGCAAGGGGGCGGGGGGGCGGAAAGCTGAGACCTTCGACCGGGAGGGAGACG	3660
Db	3601	GTGGCAAGGGGGCGGGGGGGCGGAAAGCTGAGACCTTCGACCGGGAGGGAGACG	3660
QY	3661	GCACACACACCGGCTACCGAGACGGCTGAGACCTTGACGCGACCGCTCTTGTGCTG	3720
Db	3661	GCACACACACCGGCTACCGAGACGGCTGAGACCTTGACGCGACCGCTCTTGTGCTG	3720
QY	3721	CTGCTGGGGCGACCGCGCGAGGAGCGGAGTGGCCCTCCCTTGACAGTGGGTTGGTT	3780
Db	3721	CTGCTGGGGCGACCGCGCGAGGAGCGGAGTGGCCCTCCCTTGACAGTGGGTTGGTT	3780
QY	3781	GTGTAAATAAAGTATTTAATTA	3804
Db	3781	GTGTAAATAAAGTATTTAATTA	3804
RESULT 3			
LOCUS	HUMSLA	3255 bp	DNA linear PRI 13-JUL-1995
DEFINITION	Human hormone-sensitive lipase (LIPe) gene, complete cds.		
ACCESSION	L11706		
VERSION	L11706.1	GI:896474	
KEYWORDS	hormone-sensitive lipase.		
SOURCE	Homo sapiens male adult DNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Langin,D., Laurell,H., Holst,L.S., Befrage,P. and Holm,C.		
TITLE	Gene organization and primary structure of human hormone-sensitive lipase: possible significance of a sequence homology with a lipase of Moraxella RA144, an antarctic bacterium		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 90 (11), 4897-4901 (1993)		
MEDLINE	93281663		
PUBMED	8506334		
REFERENCE	2 (bases 1 to 3255)		
AUTHORS	Holm,C.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-MAR-1993) Cecilia Holm, Cell and Molecular Biology, University of Lund, P.O. Box 94, Lund, S-221 00, Sweden		
COMMENT	On Jul 13, 1995 this sequence version replaced gi:895926.		
FEATURES	Location/Qualifiers		
source	1. 3255		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="19q13.1-q13.2"		
	/sex="male"		
	/cell_type="lymphocyte"		
	/dev_stage="adult"		
	1. 1147		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		
	/note="G00-120-154"		
	/number=1		
	1. 631		
	/gene="LIPe"		
	/note="G00-120-154"		
	join(632..1147,1148..1238,1239..1384,1385..1570,1571..1865,1866..2093,2094..2270,2271..2695,2696..2959)		
gene			
5'UTR			
exon			
3'UTR			
polyA.signal			
BASE COUNT	610 a 1057 c 1015 g 573 t		
ORIGIN			
Query Match	69.7%, Score 2650.4; DB 9; Length 3255;		
CDS			
	join(632..1147,1148..1238,1239..1384,1385..1570,1571..1865,1866..2093,2094..2270,2271..2695,2696..2959)		
	/gene="LIPe"		
	/codon_start=1		
	/product="hormone-sensitive lipase"		
	/protein_id="AA659810.1"		
	/db_xref="GI:896476"		
	/db_xref="GDB:G00-120-154"		
	/translation="MDLRTPQSLVTLAEADINAFSSSGPGETAQRISGVAFVAREDA LGLEPAUGRLGVAHLEDLPETPANQTRSLVHTARCLLAHLHKSRVTSNRSRLEF RTSHNLAEAVLYALQRLALVYAAQLVTPNGLVFEFGEGGLADLRKYVLH KSCFYGRCLGFQFPALRPFLOTISIGLVSGFKNEFGLSVASLSLTSRFLID PELRGAEPERTIOMIDVFMKAFNITEMEVLSTLAMAATVSVSLSLPEAEAM PTLADPRLTVTISPDLAHTGGCPVLVRLISDLREGDSESLIKSNORSLTELP RPQAPRSRSLIYFHGGGVPAAQTSRSHPTLRKSNMDELAPITISIDYSLAPAPRR ALBECFAYCWAIRKCALDSTGERICLAGDSAGNLCFTVALNAAVGVKRPDGLVA AYPATMDLPASPSRLSLMDPLPLVLSKCVSAVGAATKEDHSNDKALMGVLV RDRTALLRDFRLGASWNLNFEISGRKSQKSEPTAEPMRVSSEALAAQGPGLG TDSLKNLTLRDLSTLGSNSETSDPTSEMSLETLSPTSDVNLPEPADAGEAEAK NPLSPMDGLGVRAAPFEGFHPRRSSGAGTOMPLYSPIYKNPMSLLALPDSMLSL PNLVHYCALDPMDDSVMLARRLRNIGOPVTLRYVEDLPHGFLITLALCRERQAAE LCVERRILVLTTPPAGAGPSGETGAAGVGGCGGNH"		
	1148..1238		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		
	/note="G00-120-154"		
	/number=2		
	1239..1384		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		
	/note="G00-120-154"		
	/number=3		
	1385..1570		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		
	/note="G00-120-154"		
	/number=4		
	join(1571..1633,1637..1865)		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		
	/note="G00-120-154"		
	/number=5		
	1866..2093		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		
	/note="G00-120-154"		
	/number=6		
	2094..2270		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		
	/note="G00-120-154"		
	/number=7		
	2271..2695		
	/gene="LIPe"		
	/product="hormone-sensitive lipase"		</



Best Local Similarity	99.8%	Pred. No. 0;
Matches 2654;	Conservative	0; Mismatches 6; Indels 0; Gaps 0

QY	1145	CACATCCAGGATATACAGCCTCAAGGCTATCATCAACAATGGAGCTGGCAATGACAG	1200
Db	596	CCCAACCTGCCACAAAGCCTCAAGGCTATCATCAACAATGGAGCTGGCAATGACAG	655
QY	1205	TCGCTGTGACTCTGGGGGAGGACAAACATAGCCTTCTTCGAGCCAGGGTCTCTGGGAA	1264
Db	656	TCGCTGTGACTCTGGGGGAGGACAAACATAGCCTTCTTCGAGCCAGGGTCTCTGGGAA	715
QY	1265	ACGGCCCAAGCGGGCTGTAGAGCGTTTTTGGCGGTGTACGGAGCAGGCGCTGGGGCTGAG	1324
Db	716	ACGGCCCAAGCGGGCTGTAGAGCGTTTTTGGCGGTGTACGGAGCAGGCGCTGGGGCTGAG	775
QY	1325	CCGGCCCTGGGGCGCGCTGGGGGTGTGGGGCACTCTCTTGTAGCCTGGACCCAAAGACCG	1384
Db	776	CCGGCCCTGGGGCGCGCTGGGGGTGTGGGGCACTCTCTTGTAGCCTGGACCCAAAGACCG	835
QY	1385	GCCAAAGGGATACCGCAGCCTAGTGCACACAGCCCGCTGCTCGCTGGCGACCTCTCTGCAC	1444
Db	836	GCCAAAGGGATACCGCAGCCTAGTGCACACAGCCCGCTGCTCGCTGGCGACCTCTCTGCAC	895
QY	1445	AAATCCGCTATGTGGCTTCCAAACCGCCGAGCATCTTCTTCGACCAACCAACACTG	1504
Db	896	AAATCCGCTATGTGGCTTCCAAACCGCCGAGCATCTTCTTCGACCAACCAACACTG	955
QY	1505	GCCGAGCTGGAGGCTACTACGTGGCGCCCTAACCAAGTCGCGCTGGTGTACATAGCG	1564
Db	956	GCCGAGCTGGAGGCTACTACGTGGCGCCCTAACCAAGTCGCGCTGGTGTACATAGCG	1015
QY	1565	CAGCGCTGCTGTATACCAATCGGCGCGGGGGTACTTCTTGTGAGGGCGACGAGGGGCTC	1624
Db	1016	CAGCGCTGCTGTATACCAATCGGCGCGGGGGTACTTCTTGTGAGGGCGACGAGGGGCTC	1075
QY	1625	ACCGCCCACTTCTCCCGGGAGTATGTACGCTGCATTAAGGATGCTTCTATGCGCGCTGC	1684
Db	1076	ACCGCCCACTTCTCCCGGGAGTATGTACGCTGCATTAAGGATGCTTCTATGCGCGCTGC	1135
QY	1685	CTGGGCTTCCAGTTTCACAGCGCTGCACATCGGCGCATTCCTCGAGACATCTCCATTTGGGCTG	1744
Db	1136	CTGGGCTTCCAGTTTCACAGCGCTGCACATCGGCGCATTCCTCGAGACATCTCCATTTGGGCTG	1195
QY	1745	GTCGCTCTTGGGGAGCACTACAAACGCAACGAGACAGGCGCTCAGTGTGGCGCCAGCTCT	1804
Db	1196	GTCGCTCTTGGGGAGCACTACAAACGCAACGAGACAGGCGCTCAGTGTGGCGCCAGCTCT	1255
QY	1805	CTCTTTCACACAGCGCCGCTTTGCCATTCGACCCCGAGCTGCGTGGGGCTGTAGTTTACAGCG	1864
Db	1256	CTCTTTCACACAGCGCGCTTTGCCATTCGACCCCGAGCTGCGTGGGGCTGTAGTTTACAGCG	1315
QY	1865	ATCACACAGAACCTGGAGAGTGTGCATTGTGAAAGACCTTGTGAAACATCTCACAGATGGAA	1924
Db	1316	ATCACACAGAACCTGGAGAGTGTGCATTGTGAAAGACCTTGTGAAACATCTCACAGATGGAA	1375
QY	1925	GTCGTAATGCTCTTGCCCAACATGTCGATGGCCACCGTAGGGTAAAGCCGCTGCTCAGC	1984
Db	1376	GTCGTAATGCTCTTGCCCAACATGTCGATGGCCACCGTAGGGTAAAGCCGCTGCTCAGC	1435
QY	1985	CTGGCACCCGAAAGCTTTGAGATGTCCATGTACGTCCGACCCCAACGCTCTCAGGCTCACATC	2044
Db	1436	CTGGCACCCGAAAGCTTTGAGATGTCCATGTACGTCCGACCCCAACGCTCTCAGGCTCACATC	1495
QY	2045	TCACCCCACTAGGCGCCACACAGGCGCGGGCGCGCTCTGTCAGAGGCTCATCTCTCATATAC	2104
Db	1496	TCACCCCACTAGGCGCCACACAGGCGCGGGCGCGCTCTGTCAGAGGCTCATCTCTCATATAC	1555
QY	2105	CTGCGTGAAGACAGGACAGTGTGAGAGCTCAGCAGCCTATATAAGTCCAAACGGCCAAACGG	2164
Db	1556	CTGCGTGAAGACAGGACAGTGTGAGAGCTCAGCAGCCTATATAAGTCCAAACGGCCAAACGG	1615
QY	2165	AGCCTGGAAGCTGTGGCGGGCGCCCAAGACAGCACCCCGCTCGGCGTCCCTGATATGTAC	2224





RESULT 1	
AAD05012	
ID	AAD05012 standard; mRNA; 3804 BP.
XX	
AC	AAD05012;
XX	
DT	17-JUL-2001 (first entry)
XX	
DE	Human mRNA encoding testicular isoform of hormone sensitive lipase
XX	
KW	Human; hormone sensitive lipase; contraceptive; antisense-therapy;
KW	inhibitor; testicular isoform; male fertility; ss.
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	Key
FT	CDS
FT	278..3508
FT	/*tag= a
FT	/product= "Human hormone sensitive lipase"
FT	/Ec_number= "3.1.1.3"
XX	
PN	WO200126664-A2.
XX	
PD	19-APR-2001.
XX	
PF	13-OCT-2000; 2000WO-CA01228.
XX	
PR	14-OCT-1999; 99CA-2286451.
XX	
PA	(HOB1-) HOPITAL SAINTE-JUSTINE.
XX	
PI	Mitchell GA, Wang SP;

xx MPI; 2001-290653/30.  
DR P-PSDB; AAE01154.  
xx  
PT Inhibiting fertility in a male animal especially humans for  
PT contraception comprises inhibiting the activity of a hormone-sensitive  
PT lipase in the animal -  
xx  
PS Claim 10; Page 19-24; 29pp; English.

CC The present sequence is a mRNA encoding the testicular isoform of human  
CC hormone-sensitive lipase (E.C. 3.1.1.3). The present invention relates  
CC to a method of inhibiting male fertility which involves inhibition of  
CC the activity of the hormone-sensitive lipase. The hormone sensitive  
CC lipase, especially the testicular isoform from humans is useful for  
CC screening compounds that modulate or inhibit male fertility. Inhibitors  
CC of hormone sensitive lipase, especially a fully defined antisense  
CC molecule are useful for inhibiting fertility in a male animal. The  
CC method provides means for inhibiting fertility in genetically modified  
CC organisms to prevent the dissemination of genetic modifications into  
CC wild-type populations. The method is also used to identify a condition  
CC of male infertility caused by hormone sensitive lipase deficiency.  
xx

Sequence 3804 BP: 828 A; 1225 C; 1091 G; 660 U; 0 other;

Query Match 100.0%; Score 3804; DB 22; Length 3804;  
Best Local Similarity 82.6%; Pred. No. 0;  
Matches 3144; Conservative 660; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTGTGATGAGTGGTACATAGCCCTCTATCTTCTTATCTCCACCAAG 60  
DB 1 CUCUCUGUAGAGAGUGCUAGGACACUAGCCCTCCUCCUUAUCCUCCCAAG 60  
QY 61 AAAGAGCACAGAGTCTTACTTGTAGTGGGCGCAGCTGTGATCGCCAACTGCGAGTG 120  
DB 61 AAAGAGCACAGAGTCTTACTTGTAGTGGGCGCAGCTGTGATCGCCAACTGCGAGTG 120  
QY 121 CCTTAAAGAGAGACACGATGATGCTAGATGAGTGAACCCAAAGAGAGTGGCATCA 180  
DB 121 CCTTAAAGAGAGACACGATGATGCTAGATGAGTGAACCCAAAGAGAGTGGCATCA 180  
QY 121 CCUUAAGAGAGAGACGAGUGAGUGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
DB 121 CCUUAAGAGAGAGACGAGUGAGUGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 181 TGAGGAATCAATGAGAGATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 240  
DB 181 TGAGGAATCAATGAGAGATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 240  
QY 181 UGAGGAATCAATGAGAGATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 240  
DB 181 UGAGGAATCAATGAGAGATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGATGA 240  
QY 241 ACCTGGAAGATCAATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCACT 300  
DB 241 ACCTGGAAGATCAATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCACT 300  
QY 301 GTCTGAGTCAAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 301 GTCTGAGTCAAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
QY 361 GCGAG 420  
DB 361 GCGAG 420  
QY 421 ACAG 480  
DB 421 ACAG 480  
QY 481 TGCTGATCCAG 540  
DB 481 TGCTGATCCAG 540  
QY 541 TGCCCAACAG 600  
DB 541 TGCCCAACAG 600  
QY 601 ACAG 660  
DB 601 ACAG 660

QY 661 GCCAGATGAG 720  
DB 661 GCCAGATGAG 720  
QY 721 AGCTCAACAG 780  
DB 721 AGCTCAACAG 780  
QY 781 GCCATCTGCCCCAGCTGAATCTTATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
DB 781 GCCATCTGCCCCAGCTGAATCTTATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
QY 841 AACGCGAG 900  
DB 841 AACGCGAG 900  
QY 901 AAACCTGAG 960  
DB 901 AAACCTGAG 960  
QY 961 GGTCCAGATCTGAGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
DB 961 GGTCCAGATCTGAGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 1021 CACGATGAG 1080  
DB 1021 CACGATGAG 1080  
QY 1081 TAAAGTATGTCAGATGATCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
DB 1081 TAAAGTATGTCAGATGATCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
QY 1141 CAGACATGAG 1200  
DB 1141 CAGACATGAG 1200  
QY 1201 ACAAGTCTGAGTGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1201 ACAAGTCTGAGTGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
QY 1261 GGAAG 1320  
DB 1261 GGAAG 1320  
QY 1321 GGAAG 1380  
DB 1321 GGAAG 1380  
QY 1381 ACAG 1440  
DB 1381 ACAG 1440  
QY 1441 GCAAG 1500  
DB 1441 GCAAG 1500  
QY 1501 CAG 1560  
DB 1501 CAG 1560  
QY 1561 GCGAG 1620  
DB 1561 GCGAG 1620  
QY 1621 GCTACAG 1680  
DB 1621 GCTACAG 1680  
QY 1681 CAG 1740  
DB 1681 CAG 1740  
QY 1741 GCTGAGTCTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
DB 1741 GCTGAGTCTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800

Db 1741 GUGUGUCCUUCGGGAGACUACAAGCAAGAGAGAGGCTUCAGUGUGCCCGCAG 1800  
 Oy 1801 CTCTCTTTCACACAGCGCGCTTTGCCATCAGCCCCGAGCTCGTGGGCTAGATTGA 1860  
 Db 1801 CUCUCUCUCCACAGCGCGCUCUUCGACGCCGAGCUCGUGGGGCUAGUUGA 1860  
 Oy 1861 GGGATCACACAGAACCTTGAGCTTCTGAAAAGCTTGTGAACTACACCGAGAT 1920  
 Db 1861 GGGAGUACACAGAACCTUGAGCUCUUGGAAAAGCTTUCGAAACUACACCGAGAU 1920  
 Oy 1921 GGAAGTCTATGCTCTTGCCCAACATGCGATCGCCACCGTGAAGGTTAAGCCGCTCT 1980  
 Db 1921 GGAAGUGUCCUUCUGGCCAACAAUGGCAUUGGCAACCGAGGAGGUAAGCCGCTCU 1980  
 Oy 1981 CAGCTGCCACCCGAAAGCCTTTGAGATGCCATGACTGCCAGCCACGCTCAGGTCAC 2040  
 Db 1981 CAGCTGCCACCCGAAAGCCTTUGAGUCCACUAGCUCGACCCACGCTCAGGTCAC 2040  
 Oy 2041 CATCTCACCCCGCTGCGCCACAGAGCCCTGCGGCTGCTGTCAGGCTCATCTCTA 2100  
 Db 2041 CAUUCACCCCGCTCAGCGCCACAGAGCCCTUGGCGCTGCTGTCAGGCTCATCTCTA 2100  
 Oy 2101 TGAAGTCTGAGAGACAGAGAGAGAGAGCTCAGAGGCTGATTAAGTCAACAGGCA 2160  
 Db 2101 UGACCTGCGUAG 2160  
 Oy 2161 AGGAGAGCTGAGAGCTGAGCGCGCGCCACAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
 Db 2161 AGGAGAGCTGAG 2220  
 Oy 2221 GAGCTTCAAG 2280  
 Db 2221 GAGCTTCAAG 2280  
 Oy 2281 GAGCTTCAAG 2340  
 Db 2281 GAGCTTCAAG 2340  
 Oy 2341 GGGCGGCTTCCCGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
 Db 2341 GGGCGGCTTCCCGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
 Oy 2401 GGGCGGCTTCCCGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
 Db 2401 GGGCGGCTTCCCGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
 Oy 2461 GAGCTTCTTCAAG 2520  
 Db 2461 GAGCTTCTTCAAG 2520  
 Oy 2521 CATCATGAG 2580  
 Db 2521 CATCATGAG 2580  
 Oy 2581 GAGCTTCTTCAAG 2640  
 Db 2581 GAGCTTCTTCAAG 2640  
 Oy 2641 TGGTGCAG 2700  
 Db 2641 TGGTGCAG 2700  
 Oy 2701 GGTGGCGGAG 2760  
 Db 2701 GGTGGCGGAG 2760  
 Oy 2761 CAACCTCTTCTGAG 2820  
 Db 2761 CAACCTCTTCTGAG 2820  
 Oy 2821 GCGGATGCGCGCGAG 2880  
 Db 2821 GCGGATGCGCGCGAG 2880

Db 2821 GCGGATGCGCGCGAG 2880  
 Oy 2881 GCGGATGCGCGCGAG 2940  
 Db 2881 GCGGATGCGCGCGAG 2940  
 Oy 2941 GTCGAGACAG 3000  
 Db 2941 GTCGAGACAG 3000  
 Oy 3001 GTTCACCTTCTTATACACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
 Db 3001 GTTCACCTTCTTATACACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
 Oy 3061 GAGGCGGAG 3120  
 Db 3061 GAGGCGGAG 3120  
 Oy 3121 AGCTTCCAG 3180  
 Db 3121 AGCTTCCAG 3180  
 Oy 3181 CTTCATGCTGCGCGCTGCGAG 3240  
 Db 3181 CTTCATGCTGCGCGCTGCGAG 3240  
 Oy 3241 CGTGGCGTGGCGGCTGAG 3300  
 Db 3241 CGTGGCGTGGCGGCTGAG 3300  
 Oy 3301 CAACCTGAG 3360  
 Db 3301 CAACCTGAG 3360  
 Oy 3361 CTAAGCGGCGCTGCGAG 3420  
 Db 3361 CTAAGCGGCGCTGCGAG 3420  
 Oy 3421 CTTGCTGCTCACTCTCTCCCGGAG 3480  
 Db 3421 CTTGCTGCTCACTCTCTCCCGGAG 3480  
 Oy 3481 AGAGGCGGCGCTGCGAG 3540  
 Db 3481 AGAGGCGGCGCTGCGAG 3540  
 Oy 3541 TCATGAATGCTTCCCGGAG 3600  
 Db 3541 TCATGAATGCTTCCCGGAG 3600  
 Oy 3601 GTGGCAAG 3660  
 Db 3601 GTGGCAAG 3660  
 Oy 3661 GUGGCAAG 3720  
 Db 3661 GUGGCAAG 3720  
 Oy 3721 CTGCTGCGGAG 3780  
 Db 3721 CTGCTGCGGAG 3780  
 Oy 3781 GTTGTAAATTAAGTATTTAATTA 3804  
 Db 3781 GTTGTAAATTAAGTATTTAATTA 3804  
 RESULT 2  
 ABRK7218  
 ID ABRK7218 standard; cdna: 3231 BP.  
 XX  
 AC ABRK7218;  
 XX

DT	24-SEP-2002	(first entry)
XX		
DE	Human lipase, hormone-sensitive (LHPE) coding sequence.	
XX		
KV	Human: lipase; hormone sensitive; LHPE; Isozyme; obesity;	
KW	male sterility; single nucleotide polymorphism; SNP;	
KM	chromosome 19q13.1-q13.2; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..3231
FT		/tag= a
FT		/product= "Human lipase, hormone sensitive (LHPE)
FT		protein"
FT	variation	replace (280,A)
FT		/tag= b
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (295,T)
FT		/tag= c
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (298,C)
FT		/tag= d
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (436,T)
FT		/tag= e
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (525,A)
FT		/tag= f
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (634,C)
FT		/tag= g
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (1210,T)
FT		/tag= h
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (1269,T)
FT		/tag= i
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (1626,T)
FT		/tag= j
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (1677,A)
FT		/tag= k
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (1831,T)
FT		/tag= l
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (1905,A)
FT		/tag= m
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (2368,A)
FT		/tag= n
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (2429,A)
FT		/tag= o
FT		/standard_name= "Single nucleotide polymorphism (SNP)"
FT		/note= "Polymorphic site (PS)"
FT	variation	replace (2453,A)
FT		/tag= p

Query Match	Best Local Similarity	84.9%;	Score 3227.8;	DB 24;	Length 3231;
Matches 3229;	Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
278	ATGAGCCAGGTTCTTAAGTCAGTGTCTAGTCAGACTGCGACACCTGTAACACACACCAGAGG	337			

Db 1 ATGAGCCAGGTTCTAAGTCACTGTCTAGGTCAAGTGGCAACCTGAAACACACAGAGG 60  
 QY 338 CCGTAAACCCCGGTAGAGCTGGGCCGAAAGACACCCATAGCCGACGAATCGAG 397  
 Db 61 CCGTAAACCCCGGTAGAGCTGGGCCGAAAGACACCCATAGCCGACGAATCGAG 120  
 QY 398 ACTCTCAGAGGATCCAAATACCAAGACAGCTGCTTCAACCAAAAGACCCCTACCCAG 457  
 Db 121 ACTCTCAGAGGATCCAAATACCAAGACAGCTGCTTCAACCAAAAGACCCCTACCCAG 180  
 QY 458 CAGAGACCCCTGACAAATGATGCTGAATCCCAAGAACCTAGAGCCCAACAAAA 517  
 Db 181 CAGAGACCCCTGACAAATGATGCTGAATCCCAAGAACCTAGAGCCCAACAAAA 240  
 QY 518 TCTGCTTCAACAAGAAATTTCTTCCGCCACAGAACCCGACACAGCAATCACTTAC 577  
 Db 241 TCTGCTTCAACAAGAAATTTCTTCCGCCACAGAACCCGACACAGCAATCACTTAC 300  
 QY 578 ATCCAAAGGGTGTGCTCACTCAACAGAGAGCTGCTCCAGAGGAGACCTGGGCTAGGA 637  
 Db 301 ATCCAAAGGGTGTGCTCACTCAACAGAGAGCTGCTCCAGAGGAGACCTGGGCTAGGA 360  
 QY 638 AAGAAATCTATACTCAACAGAGAGCCAGCATTTGAGACAAAGACATTTAGCCCAAGG 697  
 Db 361 AAGAAATCTATACTCAACAGAGAGCCAGCATTTGAGACAAAGACATTTAGCCCAAGG 420  
 QY 698 CCGGGGCCAGAGAGAGCCACTCCAGCTCAACAGAGCTCAATACACCTGGGCCAG 757  
 Db 421 CCGGGGCCAGAGAGAGCCACTCCAGCTCAACAGAGCTCAATACACCTGGGCCAG 480  
 QY 758 GCTAAACCTGAGAGCAAAAGAGAGCCATCTGCCCGACTGAATCTACATCCCAAGAGACA 817  
 Db 481 GCTAAACCTGAGAGCAAAAGAGAGCCATCTGCCCGACTGAATCTACATCCCAAGAGACA 540  
 QY 818 CCTGAACAGTCAAGCAAGCAAGCCAGTCCAGAGGAGCCAAATCCAAAGCAAGGATCT 877  
 Db 541 CCTGAACAGTCAAGCAAGCAAGCCAGTCCAGAGGAGCCAAATCCAAAGCAAGGATCT 600  
 QY 878 TTGACAGAGCTGGGATTTCTTAACAAAATCTCAGAGACTATCCATACGCCATCA 937  
 Db 601 TTGACAGAGCTGGGATTTCTTAACAAAATCTCAGAGACTATCCATACGCCATCA 660  
 QY 938 GAGTGAAGAGCACTTCTGAGTGGTGCAGATTTCTGAGTCAGATAGATGTTGGATCA 997  
 Db 661 GAGTGAAGAGCACTTCTGAGTGGTGCAGATTTCTGAGTCAGATAGATGTTGGATCA 720  
 QY 998 TCTTCAGACAGATTTCTCAGGCCAGAGTGGTGAATGGTGGCCAGGAGTGAACCTA 1057  
 Db 721 TCTTCAGACAGATTTCTCAGGCCAGAGTGGTGAATGGTGGCCAGGAGTGAACCTA 780  
 QY 1058 GGGTAAAGGAAATCTGTTTAAAGTATGTCAAGATACAGTGGGAGCTGGCCAT 1117  
 Db 781 GGGTAAAGGAAATCTGTTTAAAGTATGTCAAGATACAGTGGGAGCTGGCCAT 840  
 QY 1118 GAGAAACAGTCTCGGATCAAGACACTACAGAGATACAGCTCAAGGCTCAATCCAC 1177  
 Db 841 GAGAAACAGTCTCGGATCAAGACACTACAGAGATACAGCTCAAGGCTCAATCCAC 900  
 QY 1178 AACATGAGCTTGGCACAATGACACAGTGGTGAATCTTGGCGAGGACAATAGCC 1237  
 Db 901 AACATGAGCTTGGCACAATGACACAGTGGTGAATCTTGGCGAGGACAATAGCC 960  
 QY 1238 TTCTTCTCAGAGGCTCTGGGGAAGAGGCCAGAGGCTGTCAAGCCGTTTTGGCGGT 1297  
 Db 961 TTCTTCTCAGAGGCTCTGGGGAAGAGGCCAGAGGCTGTCAAGCCGTTTTGGCGGT 1020  
 QY 1298 GTACGGAGCAGAGCGCTGGGCGTGAAGCCGCTGGGCGCTGCTGGGTGTGGCCAC 1357  
 Db 1021 GTACGGAGCAGAGCGCTGGGCGTGAAGCCGCTGGGCGCTGCTGGGTGTGGCCAC 1080  
 QY 1358 CTCTTTGACCTGGAGCCAGAGACCCGGCCAAACGGGTACCGGACCTAAGTGAACAGCC 1417

Db 1081 CTCTTTGACCTGGAGCCAGAGACACCGGCCAAAGGGTACCCGAGCTAGTCACACAGCC 1140  
 QY 1418 CGGTGCTGCTGGGACCTCTGACAAATCCGCTTGTGGCTTCCAAACCGCCGAGC 1477  
 Db 1141 CGGTGCTGCTGGGACCTCTGACAAATCCGCTTGTGGCTTCCAAACCGCCGAGC 1200  
 QY 1478 ATCTTCTTCCGACAGCACAACCTGGCGAGCTGGAGGCTCACTGGTGCCTCAAC 1537  
 Db 1201 ATCTTCTTCCGACAGCACAACCTGGCGAGCTGGAGGCTCACTGGTGCCTCAAC 1260  
 QY 1538 CAGCTCGCGCTGTGCTACTAGCCAGCGCTGCTGTGTACCAATGGCCGGGGTA 1597  
 Db 1261 CAGCTCGCGCTGTGCTACTAGCCAGCGCTGCTGTGTACCAATGGCCGGGGTA 1320  
 QY 1598 CTCTTCTTGGAGGCGAGAGAGGGCTACCGCCGACTTCTCCGGGATATGTACGCTG 1657  
 Db 1321 CTCTTCTTGGAGGCGAGAGAGGGCTACCGCCGACTTCTCCGGGATATGTACGCTG 1380  
 QY 1658 CATTAAGGATCTTCTATGGCCGCTGGGCTTCCAGTTCAGAGCTGGCATCCGGCCA 1717  
 Db 1381 CATTAAGGATCTTCTATGGCCGCTGGGCTTCCAGTTCAGAGCTGGCATCCGGCCA 1440  
 QY 1718 TTCTGACAGCCATCTCCATTTGGCTGTCTTCCGGGAGACACTACAAACGACAG 1777  
 Db 1441 TTCTGACAGCCATCTCCATTTGGCTGTCTTCCGGGAGACACTACAAACGACAG 1500  
 QY 1778 ACAGGCTCAATGTGGCGCCAGCTCTCTTACACAGCGCCGCTTGGCATGAGCC 1837  
 Db 1501 ACAGGCTCAATGTGGCGCCAGCTCTCTTACACAGCGCCGCTTGGCATGAGCC 1560  
 QY 1838 GAGCTGGGCGGTGAATTTGAGCGGATCAACAGAACTGGAGCTGACCTTGGGAAA 1897  
 Db 1561 GAGCTGGGCGGTGAATTTGAGCGGATCAACAGAACTGGAGCTGACCTTGGGAAA 1620  
 QY 1898 GCTTCTGGAACATCACCGAGATGAAAGTGTATGCTCTGCGCCAAACATGGATCGCC 1957  
 Db 1621 GCTTCTGGAACATCACCGAGATGAAAGTGTATGCTCTGCGCCAAACATGGATCGCC 1680  
 QY 1958 ACCGTGAGGGTAAAGCCGCTCTGAGCTGGCACCCGAAAGCTTGTGAATGCCACTGACT 2017  
 Db 1681 ACCGTGAGGGTAAAGCCGCTCTGAGCTGGCACCCGAAAGCTTGTGAATGCCACTGACT 1740  
 QY 2018 GCCGACCCAGCTCAGCGGTACCATCTCACCCCACTGGGCCACACAGGCCCTGGGCC 2077  
 Db 1741 GCCGACCCAGCTCAGCGGTACCATCTCACCCCACTGGGCCACACAGGCCCTGGGCC 1800  
 QY 2078 GTCTCTGAGGGTCAATCTCTATGACTGTGCTGAAGAGACAGCATGTAGAGCTGAGC 2137  
 Db 1801 GTCTCTGAGGGTCAATCTCTATGACTGTGCTGAAGAGACAGCATGTAGAGCTGAGC 1860  
 QY 2138 AGCTTGATTAAGTCCACAGGCCCAAGAGCTGTGAAGCTGTGGCCGCCCCACAGAGCA 2197  
 Db 1861 AGCTTGATTAAGTCCACAGGCCCAAGAGCTGTGAAGCTGTGGCCGCCCCACAGAGCA 1920  
 QY 2198 CCGCGCTCGCGGTCTGATATGTCACATTCACAGCGGTGTGTTGTGGCCAGACTGCC 2257  
 Db 1921 CCGCGCTCGCGGTCTGATATGTCACATTCACAGCGGTGTGTTGTGGCCAGACTGCC 1980  
 QY 2258 AGATCCACGAGCCCTACCTCAAGAGCTGGGCCAGAGACTGGGCGGCCCATCATCTCC 2317  
 Db 1981 AGATCCACGAGCCCTACCTCAAGAGCTGGGCCAGAGACTGGGCGGCCCATCATCTCC 2040  
 QY 2318 ATGACACTACCTCGGCGCCCTGAGGCCCTTCCCGCTGGGTGGAGAGATGCTTCTTC 2377  
 Db 2041 ATGACACTACCTCGGCGCCCTGAGGCCCTTCCCGCTGGGTGGAGAGATGCTTCTTC 2100  
 QY 2378 GCTACTGCTGGGCGCATCAAGCACTGGCCCTCTTGGCTCAAGAGGGAAGCATGTGC 2437  
 Db 2101 GCTACTGCTGGGCGCATCAAGCACTGGCCCTCTTGGCTCAAGAGGGAAGCATGTGC 2160  
 QY 2438 CTTCGGGGGACAGTGAAGCGGGAACTCTGTCTTACCTGCTCTTGGGGACAGAGCC 2497  
 Db 2161 CTTCGGGGGACAGTGAAGCGGGAACTCTGTCTTACCTGCTCTTGGGGACAGAGCC 2220



```
FT /note= "Polymorphic site 11 (PS11)"
FT 22811..22902
FT /*tag= g
FT /number= 3
FT 22903..23012
FT /*tag= r
FT /number= 3
FT 23013..23156
FT /*tag= s
FT /number= 4
FT replace (23127,T)
FT /*tag= t
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 12 (PS12)"
FT 23157..23340
FT /*tag= u
FT /number= 4
FT 23341..23527
FT /*tag= v
FT /number= 5
FT replace (23362,A)
FT /*tag= w
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 13 (PS13)"
FT replace (23516,T)
FT /*tag= x
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 14 (PS14)"
FT 23528..23664
FT /*tag= y
FT /number= 5
FT replace (23608,A)
FT /*tag= z
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 15 (PS15)"
FT 23665..23959
FT /*tag= aa
FT /number= 6
FT replace (23727,A)
FT /*tag= ab
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 16 (PS16)"
FT 23960..24744
FT /*tag= ac
FT /number= 6
FT replace (24064,G)
FT /*tag= ad
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 17 (PS17)"
FT 24745..24972
FT /*tag= ae
FT /number= 7
FT 24973..25571
FT /*tag= af
FT /number= 7
FT 25572..25748
FT /*tag= ag
FT /number= 8
FT replace (25574,A)
FT /*tag= ah
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 18 (PS18)"
FT replace (25635,A)
FT /*tag= ai
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 19 (PS19)"
FT replace (25659,A)
FT /*tag= aj
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 20 (PS20)"
FT replace (25714,T)
FT /*tag= ak
FT /standard_name= "Single nucleotide polymorphism (SNP)"
```

```
FT /note= "Polymorphic site 21 (PS21)"
FT 25749..28101
FT /*tag= al
FT /number= 8
FT 28102..28526
FT /*tag= am
FT /number= 9
FT replace (28371,A)
FT /*tag= an
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 22 (PS22)"
FT 28527..29057
FT /*tag= ao
FT /number= 9
FT replace (28958,C)
FT /*tag= ap
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 23 (PS23)"
FT replace (29012,C)
FT /*tag= ap
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 24 (PS24)"
FT replace (29014,G)
FT /*tag= ar
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 25 (PS25)"
FT replace (29018,A)
FT /*tag= as
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 26 (PS26)"
FT 29058..29321
FT /*tag= at
FT /number= 10
FT replace (29090,C)
FT /*tag= au
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 27 (PS27)"
FT replace (29261,T)
FT /*tag= av
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 28 (PS28)"
FT replace (29291,A)
FT /*tag= aw
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 29 (PS29)"
FT replace (29303,A)
FT /*tag= ax
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 30 (PS30)"
FT replace (29640,C)
FT /*tag= ay
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /note= "Polymorphic site 31 (PS31)"

Query Match 30.3%; Score 1152.8; DB 24; Length 30709;
Best Local Similarity 99.2%; Pred. No. 2.1e-239;
Matches 1151; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CTTCTTGTAGAGAGTGTAGGACATACCCCTCTATTCTATCTATCTCCACCAAG 60
DB 3807 CTTCTTGTAGAGAGTGTAGGACATACCCCTCTATTCTATCTATCTCCACCAAG 3866
QY 61 AAAGAGGCACAGAGTTCATTAGTGTGGGGCCAGCTGTATGCGCAATCGCAGCTG 120
DB 3867 AAAGAGGCACAGAGTTCATTAGTGTGGGGCCAGCTGTATGCGCAATCGCAGCTG 3926
QY 121 CCTTAAAGAGAGCCAGTGTAGTGTAGAGTGAACCCAGAGAGAGTGCATCA 180
DB 3927 CCTTAAAGAGAGCCAGTGTAGTGTAGAGTGAATGAACCCAGAGAGAGTGCATCA 3986
QY 181 TGAGGAATCAATGAGAGATCTGTGAAGAGAGAGGGCTGGTGGAGCCAGAGAGATAGA 240
DB 3987 TGAGGAATCAATGAGAGATCTGTGAAGAGAGAGGGCTGGTGGAGCCAGAGAGATAGA 4046
```



```

QY 241 ACCTGGAAGATCAATATCTCCGCTGAGGGAATTAACATGAGCCAGGTTCTAAGTCACT 300
    |||
    |||
    |||
Db 4047 ACCTGGAAGATCAATATCTCCGCTGAGGGAATTAACATGAGCCAGGTTCTAAGTCACT 4106
QY 301 GCTCTGGTCAGACTGGCAACCTGAACCAACACAGAGGCTTAATACCCCTAGAGCTTG 360
    |||
    |||
    |||
Db 4107 GCTCTGGTCAGACTGGCAACCTGAACCAACACAGAGGCTTAATACCCCTAGAGCTTG 4166
QY 361 GCCAGAAAAAGACCCCATAGCCACAGCAATGAAAGCTGTGAGGATTCATACCA 420
    |||
    |||
    |||
Db 4167 GCCAGAAAAAGACCCCATAGCCACAGCAATGAAAGCTGTGAGGATTCATACCA 4226
QY 421 ACAGAGCCCTGCTTCAAAACCAAGACCCCTCACAGAGAGACCCCTGCACCAACATGA 480
    |||
    |||
    |||
Db 4227 ACAGAGCCCTGCTTCAAAACCAAGACCCCTCACAGAGAGACCCCTGCACCAACATGA 4286
QY 481 TGCTGAATCCCAAGAGAACTAGAGCCCAACAAAAATCTGCTTCAACAAGAGAAATTTCT 540
    |||
    |||
    |||
Db 4287 TGCTGAATCCCAAGAGAACTAGAGCCCAACAAAAATCTGCTTCAACAAGAGAAATTTCT 4346
QY 541 TGCCCCACAGAAAGCCGACACAGCAATCACTTACATCCAAAGGGTGTGCTCACTCA 600
    |||
    |||
    |||
Db 4347 TGCCCCACAGAAAGCCGACACAGCAATCACTTACATCCAAAGGGTGTGCTCACTCA 4406
QY 601 ACAGAGAGCTGCTCCACAGAGGAGCTGGGCTAGAAAAAGAAATCTATTAACCAACAGA 660
    |||
    |||
    |||
Db 4407 ACAGAGAGCTGCTCCACAGAGGAGCTGGGCTAGAAAAAGAAATCTATTAACCAACAGA 4466
QY 661 GCCAGCATTTGAGACAAAACATGTAGCCCAAGGCTGGGCTGAGAGAGACCACTTC 720
    |||
    |||
    |||
Db 4467 GCCAGCATTTGAGACAAAACATGTAGCCCAAGGCTGGGCTGAGAGAGACCACTTC 4526
QY 721 AGCTCAACAAAGAGTGAATCAACACTGGGCGCCAGGCTTAACCTGAGCCAAAAGGA 780
    |||
    |||
    |||
Db 4527 AGCTCAACAAAGAGTGAATCAACACTGGGCGCCAGGCTTAACCTGAGCCAAAAGGA 4586
QY 781 GCCATCTGCCCCGAGTGAATCTACATCCCAAGAGACACTGAACAGTCAGACAAAC 840
    |||
    |||
    |||
Db 4587 GCCATCTGCCCCGAGTGAATCTACATCCCAAGAGACACTGAACAGTCAGACAAAC 4646
QY 841 AAGCCAGTCCAGGAGGCAATCCAAAGCAAGATCTTTCAGACAGCTGGGATTTCTAAC 900
    |||
    |||
    |||
Db 4647 AAGCCAGTCCAGGAGGCAATCCAAAGCAAGATCTTTCAGACAGCTGGGATTTCTAAC 4706
QY 901 AAAACTTCAGAGACTATCATACAGCGATCAAGCCCTAGAGTGAAGGACCTTTCGAGTG 960
    |||
    |||
    |||
Db 4707 AAAACTTCAGAGACTATCATACAGCGATCAAGCCCTAGAGTGAAGGACCTTTCGAGTG 4766
QY 961 GGTGCGAGATTCTGAGTCAGAAATCAGATGTGGGATCATCTTCAGACAGATTTCTCAGC 1020
    |||
    |||
    |||
Db 4767 GGTGCGAGATTCTGAGTCAGAAATCAGATGTGGGATCATCTTCAGACAGATTTCTCAGC 4826
QY 1021 CAGCATGGTGGAAATGGTGGCCAGGGAATGAAGCTTCAAGAAAAATCTGGTTA 1080
    |||
    |||
    |||
Db 4827 CAGCATGGTGGAAATGGTGGCCAGGGAATGAAGCTTCAAGAAAAATCTGGTTA 4886
QY 1081 TAAAGTGTATGTCAGAGATCACTGGGAGCTCCGACATGAGAAAAACCACTGCTCGGAATCA 1140
    |||
    |||
    |||
Db 4887 TAAAGTGTATGTCAGAGATCACTGGGAGCTCCGACATGAGAAAAACCACTGCTCGGAATCA 4946
QY 1141 CAGACACTACAGAGATACAG 1160
    |||
    |||
    |||
Db 4947 CAGACACTACAGAGATACAG 4966

```

```

RESULT 4
ABQ25066
ID ABQ25066 standard; DNA: 709 BP.
XX
AC ABQ25066;
XX
DT 12-JUL-2002 (first entry)

```

```

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 11657.
XX XX Human: cytosine methylation; 5'-CPG-3'; uracil; cytosine; diagnosis;
XX KM drug; side effect; cancer; central nervous system; cardiovascular;
XX KM gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KM SNP; cell differentiation; ds.
OS Homo sapiens.
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EPI0074.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A., Piepenbrock C., Berlin K., Guetig D;
XX WPI: 2002-371829/40.
XX
PR Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
XX amplicons from chemically treated DNA
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX cytosine methylation of a particular cytosine in a motif 5'-CPG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one
XX member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
XX and the degree of hybridisation to both classes is determined from the
XX label on the amplicon. From the ratio of labels hybridised to the two
XX classes of oligomers, the degree of methylation is calculated. The method
XX is used: (i) for diagnosis and/or prognosis of side effects of
XX therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
XX of the central nervous, cardiovascular, gastrointestinal and respiratory
XX systems etc., particularly by detecting mutations or single nucleotide
XX polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
XX types and for investigating cell differentiation. The method allows the
XX methylation status of many C residues to be determined simultaneously.
XX ABQ3410-ABQ3412 represent genomic DNA sequences used to illustrate the
XX method for determining the degree of cytosine methylation described in
XX the disclosure of the invention.
XX
SQ Sequence 709 BP; 97 A; 74 C; 238 G; 300 T; 0 other:
XX
XX Query Match 9.2%; Score 348.4; DB 24; Length 709;
XX Best Local Similarity 79.5%; Pred. No. 8.9e-66;
XX Matches 450; Conservative 0; Mismatches 111; Indels 5; Gaps 3;
QY 3244 GCGGTGCGCGTGGAGCCCAACCTGAGACAGCTGACATGCTGCGCGCGAGCGGCA 3303
    |||
    |||
    |||
Db 119 GCGGTGCGCGTGGAGTATTTATGTTGAGCAATTCGGTTATGTTGCGCGCGAGTTG 178
QY 3304 CTTGAGCCAGCGCGTGGAGCGCTGGTGGAGAGCTGCGCGAGCGCTTCTGACCT 3363
    |||
    |||
    |||
Db 179 TTTGGTTAGTGTGAGAGCTTGGCGGTGGTGGAGAGATTTGCTGACGTTTATTTT 238
QY 3364 AGCGGCTGTGCGCGAGAGCGCGCGAGCGCGAGAGCTGTGCGTGAAGCGCATCGCT 3423
    |||
    |||
    |||
Db 239 AGCGGCTGTGCGCGAGAGCGCGGTAGGTGAGTTGTGCGTGAAGCGCATTCGTTT 298
QY 3424 GGTCTCACTCTCCCGCGGAGCGCGCGAGCGCGAGAGCGGCGCTGCGGGGCTAGA 3483
    |||
    |||
    |||
Db 299 GGTCTCACTCTCTCCCGCGGAGCGCGCGGTGAGTGTGAGTGTGCGGGGCTAGA 358

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 07:06:16 : Search time 499 Seconds

(without alignments)  
17167.538 Million cell updates/sec

Title: US-09-915-814-3

Perfect score: 3804  
Sequence: 1 cctctgtaagaagagtgcta.....taataaaagatttaatta 3804

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	1.6	60	24	ABN32626
2	51	1.3	51	22	AL33801
3	51	1.3	51	22	AL33804
4	50.6	1.3	65	24	ABN28875
5	39	1.0	50	22	AL33802
6	39	1.0	50	22	AL33803
7	37.2	1.0	65	24	ABN53647
8	30.8	0.8	65	24	ABN28121
9	29.2	0.8	65	24	ABN52192

Result No.	Score	Query Match	Length	DB ID	Description
10	29	0.8	88	23	AA548737
11	27.8	0.7	89	19	AAV68382
12	27.6	0.7	60	24	ABN45730
13	27	0.7	60	24	ABN32825
14	27	0.7	72	20	AAV88105
15	26.8	0.7	99	17	AAV27118
16	26.8	0.7	99	18	AAV96968
17	26.6	0.7	65	21	AAV88283
18	26.6	0.7	65	22	AAV74335
19	26.6	0.7	65	22	AAV23371
20	26.4	0.7	100	22	AAV45837
21	26.4	0.7	100	22	AAV45837
22	26.4	0.7	100	24	ABN20120
23	26.2	0.7	86	24	AAV98931
24	26.2	0.7	87	16	AAV22575
25	26	0.7	90	24	AAV36582
26	26	0.7	92	21	AAV11989
27	26	0.7	94	19	AAV38341
28	25.8	0.7	99	21	AAV36736
29	25.6	0.7	64	20	AAV88104
30	25.6	0.7	65	24	AAV57988
31	25.6	0.7	66	12	AAQ12354
32	25.6	0.7	70	21	AAV38176
33	25.6	0.7	99	14	AAQ36997
34	25.6	0.7	99	21	AAV98715
35	25.6	0.7	100	24	ABV74850
36	25.4	0.7	96	21	AAV57144
37	25.4	0.7	100	17	AAV15244
38	25.2	0.7	61	18	AAV84492
39	25.2	0.7	61	21	AAV82882
40	25.2	0.7	61	22	AAV74334
41	25.2	0.7	61	22	AAV23370
42	25.2	0.7	65	19	AAV41235
43	25.2	0.7	68	22	AAV32956
44	25.2	0.7	100	22	AAV41935
45	25	0.7	66	17	AAV97798

#### ALIGNMENTS

RESULT 1  
ABN32626  
ID ABN32626 standard; DNA: 60 BP.  
XX  
AC ABN32626;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human spliced transcript detection oligonucleotide SEQ ID NO:5374.  
XX  
KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 20-JUL-2001; 2001WO-1B01903.  
XX  
PR 28-JUL-2000; 2000US-221607P.  
XX  
PR 02-MAY-2001; 2001US-287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Masserman A, Mintz E, Mintz L, Fatigler S;  
XX  
DR WPI; 2002-257383/30.  
XX  
PT New oligonucleotide libraries comprising oligonucleotides which selectively hybridize to mRNAs transcribed from a transcription unit of

PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

PS Example 1: SEQ ID 5374; 47pp; English.

CC The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the  
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises  
CC several oligonucleotides, each capable of hybridizing selectively to a  
CC set of messenger RNAs transcribed from a given transcription unit of  
CC the genome, which encodes one or more messenger RNA splice variants.  
CC The oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterizing the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcripts. The libraries may also be used as specialized mini  
CC libraries to detect transcripts of a sub-transcriptome under a  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition; to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN55589 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in  
CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 60 BP; 19 A; 16 C; 16 G; 9 T; 0 other;

Query Match 1.6%; Score 60; DB 24; Length 60;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2621 AAGTGTGTCAGCGCTATGCTGTGCAAGAGAGAGACACACACACACACAGAA 2680  
DB 1 AAGTGTGTCAGCGCTATGCTGTGCAAGAGAGAGACACACACACACAGAA 60

RESULT 2

AA133801

ID AA133801 standard; DNA; 51 BP.

AC AA133801;

DT 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #7009.

KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.

KW Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

PA (CURA-) CURAGEN CORP.

XX

PI Shinkets RA, Leach M;

DR WPI; 2001-465210/50.

PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g.  
PT cancer, autoimmune diseases and infections

PS Claim 1: Page 3391; 4143pp; English.

CC The present invention relates to oligonucleotides encoding polymorphic  
CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
CC histones, kinases, colony stimulating factors, complement related  
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
CC G-protein coupled receptors and thioesterases. The present sequence is  
CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
CC by them may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate expression of the proteins listed  
CC above. Disorders that may be prevented, diagnosed and/or treated include  
CC multifactorial diseases with a genetic component, such as autoimmune  
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
CC leukaemia), diseases of the nervous system and an infection of pathogenic  
CC organisms.

SQ Sequence 51 BP; 6 A; 24 C; 11 G; 10 T; 0 other;

Query Match 1.3%; Score 51; DB 22; Length 51;

Best Local Similarity 100.0%; Pred. No. 0.0099;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1516 GGCCTACCTGGCTGCTCCACACAGCTCCGCTGCTACTACGCCCA 1566

DB 1 GGCCTACCTGGCTGCTCCACACAGCTCCGCTGCTACTACGCCCA 51

RESULT 3

AA133804

ID AA133804 standard; DNA; 51 BP.

AC AA133804;

DT 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #7012.

KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease; ss.

KW Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

PA (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

DR WPI; 2001-465210/50.

XX polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 PT cancer, autoimmune diseases and infections -  
 PS Claim 1; Page 3392; 4143pp; English.

XX The present invention relates to oligonucleotides encoding polymorphic  
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 CC histones, kinases, colony stimulating factors, complement related  
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
 CC G-protein coupled receptors and thioesterases. The present sequence is  
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
 CC by them may be used in the prevention, diagnosis and treatment of  
 CC diseases associated with inappropriate expression of the proteins listed  
 CC above. Disorders that may be prevented, diagnosed and/or treated include  
 CC multifactorial diseases with a genetic component, such as autoimmune  
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
 CC leukemia), diseases of the nervous system and an infection of pathogenic  
 CC organisms.

XX Sequence 51 BP; 7 A; 21 C; 18 G; 5 T; 0 other;

Query Match 1.3%; Score 51; DB 22; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 0.099;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2160 AACGAGACTGTGAGCTGTGCGCCGCCAGCAGCAGCCCGTCGCGGT 2210  
 DB 1 AACGAGACTGTGAGCTGTGCGCCGCCAGCAGCAGCCCGTCGCGGT 51

RESULT 4  
 ABN28875  
 ID ABN28875 standard; DNA; 65 BP.

XX ABN28875;

XX 15-JUL-2002 (first entry)

DE Rat spliced transcript detection oligonucleotide SEQ ID NO:1623.

XX Human; mouse; rat; splice transcript; detection; RNA transcript;  
 KW splice variant; transcriptome; oligonucleotide library; ss.

XX Rattus norvegicus.

XX WO200210449-A2.

XX 07-FEB-2002.

XX 20-JUL-2001; 2001WO-1B01903.

XX 28-JUL-2000; 2000US-221607P.

XX 02-MAY-2001; 2001US-287724P.

XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of  
 PT a genome, useful for detecting tissue-, pathology-, and  
 PT developmental-specific genes -

XX Example 1; SEQ ID 1623; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting

CC messenger RNAs that populate a (sub-)transcriptome, where the  
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
 CC transcription units that populate a genome. The library comprises  
 CC several oligonucleotides, each capable of hybridising selectively to a  
 CC set of messenger RNAs transcribed from a given transcription unit of  
 CC the genome, which encodes one or more messenger RNA splice variants.  
 CC The oligonucleotide libraries are useful for detecting mRNAs from a  
 CC biological sample, in expression profiling studies, in qualitatively or  
 CC quantitatively characterising the corresponding transcriptome, and in  
 CC detecting RNA transcripts and splice variants of human or animal  
 CC transcriptomes. The libraries may also be used as specialised mini  
 CC libraries to detect transcripts of a sub-transcriptome under a  
 CC particular biological or pathological state, and so allowing the  
 CC detection of tissue- and pathology-specific genes such as those genes  
 CC only expressed in specific tissue under a specific pathological  
 CC condition; to detect developmental specific genes; and to detect RNA  
 CC transcripts and splice variants of a transcriptome of a patient suffering  
 CC from a particular disorder. ABN27253 to ABN59589 represent  
 CC oligonucleotide sequences from rats, humans and mice, which are used in  
 CC the exemplification of the present invention.  
 CC N.B. The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 65 BP; 13 A; 26 C; 9 G; 17 T; 0 other;

Query Match 1.3%; Score 50.6; DB 24; Length 65;  
 Best Local Similarity 86.2%; Pred. No. 0.13;  
 Matches 56; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3132 AGGTGCCACAGATGCGCCCTCTACTCTACCCATGATGCAAGAACCCCTTCATGCGC 3191  
 DB 1 AAGGTGCTCTCCTCAGATGCGCCCTCTACTCTGTCACCATGATGCAAGAACCCCTTCATGCTGC 60

QY 3192 CGCTG 3196

DB 61 CTCTG 65

RESULT 5  
 AAL3802  
 ID AAL3802 standard; DNA; 50 BP.

XX AAL3802;

XX 24-JAN-2002 (first entry)

DE Human SNP oligonucleotide #7010.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;  
 KW amyloid protein; angiotensin, apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW multifactorial disease; G-protein coupled receptor; thioesterase; inflammation;  
 KW nervous system disease; infection;

XX Homo sapiens.

XX WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

XX 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

DR WPI; 2001-465210/50.  
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
 PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 PT cancer, autoimmune diseases and infections -  
 PS Claim 1; Page 3392; 4143pp; English.  
 XX  
 CC The present invention relates to oligonucleotides encoding polymorphic  
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 CC histones, kinases, colony stimulating factors, complement related  
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
 CC G-protein coupled receptors and thioesterases. The present sequence is  
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
 CC by them may be used in the prevention, diagnosis and treatment of  
 CC diseases associated with inappropriate expression of the proteins listed  
 CC above. Disorders that may be prevented, diagnosed and/or treated include  
 CC multifactorial diseases with a genetic component, such as autoimmune  
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
 CC leukaemia), diseases of the nervous system and an infection of pathogenic  
 CC organisms.  
 CC  
 SQ Sequence 50 BP; 7 A; 16 C; 15 G; 12 T; 0 other;  
 XX  
 Query Match 1.0%; Score 39; DB 22; Length 50;  
 Best Local Similarity 98.0%; Pred. No. 39;  
 Matches 50; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1596 TACTCTTCTTGTGAGGCGAGCGAGGCTCACCCTCTCTCCGAGGT 1646  
 Db 1 TACTCTTCTTGTGAGGCGAGCGAGGCTCACCCTCTCTCCGAGGT 50  
 ||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 6  
 AAL33803  
 ID AAL33803 standard; DNA; 50 BP.  
 XX  
 AC AAL33803;  
 XX  
 DT 24-JAN-2002 (first entry)  
 XX  
 DE Human SNP oligonucleotide #7011.  
 XX  
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
 KW neuroprotective; antimitochondrial; gene therapy; vaccine; amylose; cancer;  
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;  
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
 KW multifactorial disease; autoimmune disease; infection;  
 KW nervous system disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200147944-A2.  
 XX  
 PD 05-JUL-2001.  
 XX  
 PF 28-DEC-2000; 2000WO-US35498.  
 XX  
 PR 28-DEC-1999; 99US-0173419.  
 PR 27-DEC-2000; 2000US-0173419.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Leach M;  
 DR WPI; 2001-465210/50.  
 XX  
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

PT oncogenes and histones, useful for diagnosing and treating, e.g.  
 PT cancer, autoimmune diseases and infections -  
 PS Claim 1; Page 3392; 4143pp; English.  
 XX  
 CC The present invention relates to oligonucleotides encoding polymorphic  
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,  
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,  
 CC histones, kinases, colony stimulating factors, complement related  
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,  
 CC G-protein coupled receptors and thioesterases. The present sequence is  
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded  
 CC by them may be used in the prevention, diagnosis and treatment of  
 CC diseases associated with inappropriate expression of the proteins listed  
 CC above. Disorders that may be prevented, diagnosed and/or treated include  
 CC multifactorial diseases with a genetic component, such as autoimmune  
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,  
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer  
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,  
 CC leukaemia), diseases of the nervous system and an infection of pathogenic  
 CC organisms.  
 CC  
 SQ Sequence 50 BP; 10 A; 16 C; 15 G; 9 T; 0 other;  
 XX  
 Query Match 1.0%; Score 39; DB 22; Length 50;  
 Best Local Similarity 98.0%; Pred. No. 39;  
 Matches 50; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1612 CGACGAGGGGCTCACCCTCTCTCCGAGGTATGTCACGGTCATTA 1662  
 Db 1 CGACGAGGGGCTCACCCTCTCTCCGAGGTATGTCACGGTCATTA 50  
 ||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 7  
 ABN53647  
 ID ABN53647 standard; DNA; 65 BP.  
 XX  
 AC ABN53647;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:26395.  
 XX  
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
 KW splice variant; transcriptome; oligonucleotide library; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200210449-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 20-JUL-2001; 2001WO-IB01903.  
 XX  
 PR 28-JUL-2000; 2000US-221607P.  
 PR 02-MAY-2001; 2001US-287724P.  
 XX  
 PA (COMP-) COMPUGEN INC.  
 XX  
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
 DR WPI; 2002-257383/30.  
 XX  
 PT New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of  
 PT a genome, useful for detecting tissue-, pathology-, and  
 PT developmental-specific genes -  
 PS Example 1; SEQ ID 26395; 47pp; English.  
 XX  
 CC The present invention describes oligonucleotide libraries for detecting  
 CC messenger RNAs that populate a (sub-)transcriptome, where the  
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2003, 04:03:01 ; Search time 3324 Seconds  
(without alignments)  
18534.185 Million cell updates/sec

Title: US-09-915-814-3  
Perfect score: 3804  
Sequence: 1 ctctctgtaagagagtgctca.....taataaaagtatttaatta 3804

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estinu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_huv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_trod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1792.2	47.1	1809	11	BC029961 Homo sapi
2	1213	31.9	2760	11	AK019846 Mus muscu
3	921	24.2	1084	13	BM563641 AGENCOURT
4	808.6	21.3	826	13	BI826568 603077016
5	787	20.7	872	13	BI827841 603074047
6	784	20.6	905	13	BI827706 603074127

Result No.	Score	Query Match	Length	DB ID	Description
7	782.6	20.6	870	13	BI827559 603073524
8	764.8	20.1	871	13	BI561175 603253712
9	757	19.9	787	13	BI830913 603081040
10	753.2	19.8	788	13	BI520246 603071104
11	726	19.1	773	13	BI520160 603071405
12	711.2	18.7	781	13	BI832284 603080125
13	697	18.3	767	13	AI337272 lb96b07.x
14	668.8	17.6	1367	11	BC029301 Homo sapi
15	652	17.1	663	13	BI826036 603076441
16	645	17.0	645	13	BI825158 603072030
17	623	16.4	625	13	BM310576 603072030
18	619	16.3	668	14	BQ028786 UI-H-DF0-
19	613.8	16.1	736	9	AI953697 wq47c09.x
20	611	16.1	691	13	BI859566 603383330
21	610	16.0	647	13	BI818150 603032384
22	568	14.9	569	10	AM269990 xv46n03.x
23	567.2	14.9	588	14	BQ807836 NISC_K10
24	558	14.7	985	12	BG023944 602303637
25	556.2	14.6	653	9	AI393944 tg11b06.x
26	555.8	14.6	601	9	AI582719 t016907.x
27	548.2	14.4	555	9	AI798896 w634a09.x
28	538.6	14.2	635	9	AA601541 m087n07.s
29	534	14.0	534	12	BE855571 7910d02.x
30	521.4	13.7	741	12	BG872368 602792655
31	520.4	13.7	524	10	AM195614 xm84g08.x
32	510.4	13.4	892	14	BQ953768 AGENCOURT
33	503.2	13.2	669	9	AA861626 ak34f10.s
34	500.6	13.2	1036	13	BM560162 AGENCOURT
35	499.6	13.1	557	14	BM714935 UI-E-EJ0-
36	498.8	13.1	563	14	BM676647 UI-E-EJ0-
37	496.6	13.1	572	9	AA639729 n982a04.s
38	496	13.0	539	10	AM027306 wt73a09.x
39	492.2	12.9	643	12	BF222934 7925c09.x
40	479	12.6	952	12	BF537918 602049357
41	477.4	12.5	882	12	BG868403 602787874
42	477.2	12.5	913	13	BI654202 603280304
43	473.2	12.4	600	12	BG803367 0224-29 M
44	473	12.4	588	12	BG835117 353551 MA
45	469	12.3	469	10	AM136527 UI-H-B11-

## ALIGNMENTS

RESULT 1  
BC029961  
LOCUS 1809 bp mRNA linear HTC 06-MAY-2002  
DEFINITION Homo sapiens, clone IMAGE:5171623, mRNA.  
ACCESSION BC029961  
VERSION BC029961.1 GI:20455827  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA  
REMARK  
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Gunaratne, P.H., Garcia, A.M., Lu, X., Huiy, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAC Plate: 51 Row: f Column: 9  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4885450  
This clone has the following problem: no polyA-tail.

## FEATURES

## Source

Location/Qualifiers  
1..1809  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5171623"  
/tissue\_type="Brain, adult medulla"  
/clone\_lib="NIH\_MGC\_119"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
BASE COUNT 475 a 541 c 480 g 313 t  
ORIGIN

Query Match 47.1%; Score 1792.2; DB 11; Length 1809;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1805; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
Oy 15 GTGTAGGACATAGACCCCTCTATTCCTATCCCTCCACCAAGAAAGAGCAGAG 74
    |||||
Db 1 GTGTAGGACATAGACCTCTCTCTATTCCTATCCCTCCACCAAGAAAGAGCAGAG 60
Oy 75 TTGATTACTAGTGGGGGCGAGCTGTGATGGCCAACTGCCAGTGCCTTAAAAAGAG 134
    |||||
Db 61 TTGATTACTAGTGGGGGCGAGCTGTGATGGCCAACTGCCAGTGCCTTAAAAAGAG 120
Oy 135 ACCAGTGTAGTGAAGAGAGTGAAGAAACCAAGAGAGTGCATATGAGGATCAATGA 194
    |||||
Db 121 ACCAGTGTAGTGAAGAGAGTGAAGAAACCAAGAGAGTGCATATGAGGATCAATGA 180
Oy 195 GAGATCTGTGAAGAGAGAGGCTGGTGGAGCCAGAGAGATAGAACTGGAGATCAA 254
    |||||
Db 181 GAGATCTGTGAAGAGAGAGGCTGGTGGAGCCAGAGAGATAGAACTGGAGATCAA 240
Oy 255 TATCTCCGCTGAGGAGAAATACAAATGAGCCAGGTTCTAAGTCACTGTAGTCAACT 314
    |||||
Db 241 TATCTCCGCTGAGGAGAAATACAAATGAGCCAGGTTCTAAGTCACTGTAGTCAACT 300
Oy 315 GGCACCTGAACACACACAGAGGCTATTAACCCGCTAGAGCTGGGCCAGAAAAGACAC 374
    |||||
Db 301 GGCACCTGAACACACACAGAGGCTATTAACCCGCTAGAGCTGGGCCAGAAAAGACAC 360
Oy 375 CCATAGCCAGCAGAAATGGAAGACTCTGCAAGGATCCAAATACCCAAAGAGCCTGCTT 434
    |||||
Db 361 CCATAGCCAGCAGAAATGGAAGACTCTGCAAGGATCCAAATACCCAAAGAGCCTGCTT 420
Oy 435 CAACCAAAAGACCCCTCACCCAGAGAGACCCCTGACACAACTGATGCTGAATCCAGA 494
    |||||
Db 421 CAACCAAAAGACCCCTCACCCAGAGAGACCCCTGACACAACTGATGCTGAATCCAGA 480
Oy 495 AGGACCTAGAGCCCAACAAATGCTTCAAGAGAAATTTCTTGGCCCAACAGAAC 554
    |||||
Db 481 AGGACCTAGAGCCCAACAAATGCTTCAAGAGAAATTTCTTGGCCCAACAGAAC 540
Oy 555 CCGACACAGACATACCTTACATCCAAAGGCTGCTGCTACTCAACAGAAAGCTGCT 614
    |||||
Db 541 CCGACACAGACATACCTTACATCCAAAGGCTGCTGCTACTCAACAGAAAGCTGCT 600
Oy 615 CCCAGCAGGAGCCTGGGCTAGGAAAAAGATCTATACTCAACAGAGACCGAGATTGAGAC 674
    |||||
Db 601 CCCAGCAGGAGCCTGGGCTAGGAAAAAGATCTATACTCAACAGAGACCGAGATTGAGAC 660
Oy 675 AAAGACATGTAGCCCAAGAGGCTGGGCCAGAGAGACCTCAGTCAACAAGAG 734
    |||||
Db 661 AAAGACATGTAGCCCAAGAGGCTGGGCCAGAGAGACCTCAGTCAACAAGAG 720
```

```
Oy 735 CTGAATCAACACCTGGGCCAGGCTAAACCT-GGAGCCCAAAAGGAGCATCTGCCCG 793
    |||||
Db 721 CTGAATCAACACCTGGGCCAGGCTAAACCTGGAGGCCAAAAGGAGCATCTGCCCG 780
Oy 794 ACTGAATCTACATCCCAAGACAGACACCTGACAGTCAAGACAAACAGCCATGCCAG 853
    |||||
Db 781 ACTGAATCTACATCCCAAGACAGACACCTGACAGTCAAGACAAACAGCCATGCCAG 840
Oy 854 GGAGCCAAATCCAGAGAGGATCTTGAACAGAGTGGATTTCTTAACAAATTCAGAA 913
    |||||
Db 841 GGAGCCAAATCCAGAGAGGATCTTGAACAGAGTGGATTTCTTAACAAATTCAGAA 900
Oy 914 CTATCCATACAGCGATCAGCCCTAGAGTGAAGGCACTTCTGAGTGGTCCAGATTCT 973
    |||||
Db 901 CTATCCATACAGCGATCAGCCCTAGAGTGAAGGCACTTCTGAGTGGTCCAGATTCT 960
Oy 974 GAGTCAGAAATCAGATGTGGATCATCTTGAACACAGATTTCTCCAGCCACGATGGGTGA 1033
    |||||
Db 961 GAGTCAGAAATCAGATGTGGATCATCTTGAACACAGATTTCTCCAGCCACGATGGGTGA 1020
Oy 1034 ATGTGGCCAGAGGAGTGAAGCTAGGCTCAAGGAAATCTGTATTAAGTGAATGA 1093
    |||||
Db 1021 ATGTGGCCAGAGGAGTGAAGCTAGGCTCAAGGAAATCTGTATTAAGTGAATGA 1080
Oy 1094 GGATACAGTGGAGCGTCCGACATGAGAAACAGTGTCTGGAAATCAGACACTACAG 1153
    |||||
Db 1081 GGATACAGTGGAGCGTCCGACATGAGAAACAGTGTCTGGAAATCAGACACTACAG 1140
Oy 1154 GATACAGCTCAAGGCTCATCCACAAGATGACCTGCGCAATATGACATAGTCGCTGGTG 1213
    |||||
Db 1141 GATACAGCTCAAGGCTCATCCACAAGATGACCTGCGCAATATGACATAGTCGCTGGTG 1200
Oy 1214 ACTCTGGCGGAGAGCAACATATAGCTTCTCTGAGGACAGGTCCTGGGAAAGCGGCCAG 1273
    |||||
Db 1201 ACTCTGGCGGAGAGCAACATATAGCTTCTCTGAGGACAGGTCCTGGGAAAGCGGCCAG 1260
Oy 1274 CGGCTGTACGGGCTTTTTCGCGGTGTACGGGAGCAGGCGCTGGGCTGAGCCGCGCTTG 1333
    |||||
Db 1261 CGGCTGTACGGGCTTTTTCGCGGTGTACGGGAGCAGGCGCTGGGCTGAGCCGCGCTTG 1320
Oy 1334 GGCGCGCTGCTGGGCTGGCGGACCTCTTGAACCTGGAGCCAGAGACACCGGCCAAAGGG 1393
    |||||
Db 1321 GGCGCGCTGCTGGGCTGGCGGACCTCTTGAACCTGGAGCCAGAGACACCGGCCAAAGGG 1380
Oy 1394 TACCGAGCCTAGTGCACACAGCCGCTGCTGCTGCGCACCTCTGCACAAATCCCGC 1453
    |||||
Db 1381 TACCGAGCCTAGTGCACACAGCCGCTGCTGCTGCGCACCTCTGCACAAATCCCGC 1440
Oy 1454 TATGTGGCTCCAAACCGCGCAGACATCTTCTCCGACACAGCCACCACTGGCCAGCTTG 1513
    |||||
Db 1441 TATGTGGCTCCAAACCGCGCAGACATCTTCTCCGACACAGCCACCACTGGCCAGCTTG 1500
Oy 1514 GAGGCTACTGCTGGCTCCACCCAGAGCTCCGCGCTGCTGCTAGTACAGCCAGAGGCTTG 1573
    |||||
Db 1501 GAGGCTACTGCTGGCTCCACCCAGAGCTCCGCGCTGCTGCTAGTACAGCCAGAGGCTTG 1560
Oy 1574 CTGGTTACCAATGCGCGGGGTACTCTTCTTGAAGGCGAGAGGGGCTCACCGCGCAG 1633
    |||||
Db 1561 CTGGTTACCAATGCGCGGGGTACTCTTCTTGAAGGCGAGAGGGGCTCACCGCGCAG 1620
Oy 1634 TTCTCCGCGGAGATATGTCAAGCTGCATAAGGATGCTTATAGCCGCTGCTGGGCTTG 1693
    |||||
Db 1621 TTCTCCGCGGAGATATGTCAAGCTGCATAAGGATGCTTATAGCCGCTGCTGGGCTTG 1680
Oy 1694 CAGTTACGCTCTGCCATCTCCGCCATTTCTGCAAGACATCTCCATTGGTGGTGTCTTC 1753
    |||||
Db 1681 CAGTTACGCTCTGCCATCTCCGCCATTTCTGCAAGACATCTCCATTGGTGGTGTCTTC 1740
Oy 1754 GGGGAGCACTAACAACGCAAGGAGAGGCTCAGTGTGGCGGCAAGCTCTCTCTTACCC 1813
    |||||
Db 1741 GGGGAGCACTAACAACGCAAGGAGAGGCTCAGTGTGGCGGCAAGCTCTCTCTTACCC 1800
```





SWINSELEISGRKPKRTSPTAGELISPKHLAKTQNTLITATLVLIVLQALSL  
 PROQTSMFCCFPROCHNSRNOIKKKKKRREGK<sup>R</sup>  
 BASE COUNT 636 a 802 c 758 g 564 t  
 ORIGIN

Query Match 31.9%; Score 1213; DB 11; Length 2760;  
 Best Local Similarity 82.5%; Pred. No. 1.3e-264;  
 Matches 1403; Conservative 0; Mismatches 295; Indels 3; Gaps 1;

QY 1119 AGAAGACAGTGTCTGGGAATACAGACACTACAGATACAGCTTAAGGCTCATCCACA 1178  
 DB 847 AGAAGACAGTGTCTGGGAATACAGACACTACAGATACAGCTTAAGGCTCATCCACA 906  
 QY 1179 ACATGACCTGCGACATACAGACACTGCTGTGACTGTGGGAGGAGCAATATGCTT 1238  
 DB 907 GCATGATTTAGCAGATGAGACACTGCTGTGACTGTGGGAGGAGCAATATGCTT 966  
 QY 1239 TCTTCTGAGCCAGAGGCTCTGGGGAAGGCGGCTGTCAAGCGTTTTCGGGTG 1298  
 DB 967 TCTTCTGAGCCAGAGGCTCTGGGGAAGGCGGCTGTCAAGCGTTTTCGGGTG 1026  
 QY 1299 TAGGAGACGAGGCTGTGGGCTGAGACCGGCGCTGGGCGGCTGGGCTGGGCTG 1358  
 DB 1027 TTGGGGAACAGGACACTGGGCTGTGAAACCAACCTTAGGCCAATGTTGGGTGGCACACC 1086  
 QY 1359 TCTTGTACCTGAGCCAGAGACACCGGCGACGAGGCTTACGAGCTTACACAGCC 1418  
 DB 1087 ATTTTGTACCTGAGCCAGAGACACCGGCGACGAGGCTTACGAGCTTACACAGCC 1146  
 QY 1419 GCTGTGCTGGGCGACCTCTGCAACAAATCCCGCTATGTGGCTTCAACCGCCGACGA 1478  
 DB 1147 GATGTGCTGGGCGACCTCTGCAACAAATCCCGCTATGTGGCTTCAACCGCCGACGA 1206  
 QY 1479 TCTTCTTCCGACAGCCAACTGTGGCGAGCTGGAGGCTTACCTGGGCTGGCTCACCC 1538  
 DB 1207 TCTTCTTCCGAGCTGTGCAACCACTGAGAGCTTGGAGGCTTACCTGGGCTGGCTCACCC 1266  
 QY 1539 AGCTCCGCGCTGTGTCTACTACGCCAGCGCTGTGTGTTACCAATCGCCGGGGGTAC 1598  
 DB 1267 AGCTCCGCGCTGTGTCTACTACGCCAGCGCTGTGTGTTACCAATCGCCGGGGGTAC 1326  
 QY 1599 TCTTCTTGGGCGAGCGAGGCTCACCGCGAGCTTCTCCGAGAGATATGTCAGAGCTGC 1658  
 DB 1327 TCTTCTTGGAGGCTGTATGAGAGACTCACCGCTGACTTCTGTGAAGATATGTCAGAGCTGC 1386  
 QY 1659 ATAGAGATGCTTCTATGAGCGGCTGCTGGGCTTCCAGTTACAGGCTTCCATCCGCCAT 1718  
 DB 1387 ACAAGGCTGTCTTACAGCGGCTGCTGGGCTTCCAGTTACAGGCTTCCATCCGCCAT 1446  
 QY 1719 TCTGCAACACCATCTCCATTTGGGCTGTGTCTTGGGAGGACTTCAAAACGCAAGAGA 1778  
 DB 1447 TCTGCAACACCATCTCTCCATGCGGGGTGTCTTGGGAGGACTTCAAAACGCAAGAGA 1506  
 QY 1779 CAGGCTCAGTGTGGCGCGAGCTCTCTTCACACAGCGGCGCTTCCATTCGACCCCG 1838  
 DB 1507 CAGGCTCAGTGTGGCGCGAGCTCTCTTCACCGGCTGGCGGATTCGCCATAGACCCAG 1566  
 QY 1839 AGCTGCGTGGGCTGTAGTTTACGGGATCACAGAAACCTTGGAGCTTCTTGAAG 1898  
 DB 1567 AGTGTGCGGGGCTGTAAATTTTAAAGCATATACAGAAACCTTGGAGCTTCTTGAAG 1626  
 QY 1899 CCTTCTGGAACATCAGAGAGATGGAAGGCTATGCTCTTGGCCAAACATGCAATCGGCCA 1958  
 DB 1627 CCTTCTGGAACATCAGTGAAGTGTGAGGCTGTCTTGGCCAAACATGCAATCGGCCA 1686  
 QY 1959 CCGTGAAGGTAAGCGGCTGTCTACGCTGACCGCGAAGCTTGTGAGTGTGCACTGCTG 2018  
 DB 1687 CTGTAAGGTAAGCGGCTGTCTACGCTGACCGCGTGTGAGGCTTGTGAGTGTGCACTGCT 1746  
 QY 2019 CCGAGCCAGCTGTAGGCTACATCTTACCGCCAGTGGCCACAGAGGCGCTGGGCGG 2078  
 DB 1747 CTGATCCAGGCTTACAGTATACATCTCCTTGGCAACACAGGAGACAGCTCTTG 1806

QY 2079 TCTCTGTCAGGCTCATCTCTTATGACCTGCTGAAGAGACAGAGAGTGAAGCTCAGCA 2138  
 DB 1807 TGCTAGCCAGGCTCATCTCTTATGACCTAGGGAAGAGACAGAGAGTATCTACACA 1866  
 QY 2139 GCTGTAAAGTCCAAAGGCGCAACGAGGCTGAGCTGTGGCGGCGCCCGCCAGAGGAC 2198  
 DB 1867 GCTGTAAAGTCCAAAGGCGCGAC--GCTGTAGCTGTGGCGGCGCCCGCCAGAGGAC 1923  
 QY 2199 CCGGCTGCGGCTCCGATAGTGCATTTCCAGCGGCGGCTTGTGGCCAGACTTCCA 2258  
 DB 1924 CCGGTTACAGGCGCGCTGTGTTCATCATCCAGAGGCGGCTTGTGGCACAGACTCTA 1983  
 QY 2259 GATCCACAGAGCTTACCTCAAGAGCTGGGCGCAGAGAGCTGGGCGCCCATCATCTCA 2318  
 DB 1984 AATCCACAGAGCTTACCTCAAGAGCTGGGCGCAGAGAGCTAGAGCTTCTCTCA 2043  
 QY 2319 TCGACTACCTCCGCGCCCTGTAGGCGCCCTTCCCGCTGCGGTGAGAGAGTCTCTCG 2378  
 DB 2044 TCGACTACCTCCGCGCCCGCGAGGCTCCCTTCCCGAGCGCTGAGAGAGTCTCTCG 2103  
 QY 2379 CCTACTGCTGGGCTATCAAGCACTGCGCCCTCTTGGCTCAACAGGGAACGAATCTGCC 2438  
 DB 2104 CCTACTGCTGGGCTGTCAAGCACTGCTGACTGTGCTTGTCAAGTGAAGAGGATATGCC 2163  
 QY 2439 TTGCGGGGACAGTGCAGAGCGGAACCTGTGTCACCGTGGCTGTGGGACAGAGCT 2498  
 DB 2164 TTGCGGGGACAGTGCAGAGCGGAACCTGTGTCACCGTGGCTGTGGGACAGAGCT 2223  
 QY 2499 ACGGGGCGGGGTGCAGATGGCATATGCAAGCTTACCGCGGCGCAATGTGTGAGCTG 2558  
 DB 2224 ATGGAAGAGGCTGCAGATGGCATATGCAAGCTTACCGAGTATACCAACCTGTGAGTCT 2283  
 QY 2559 CCGGCTCTCCCTCCCGGCTGCGAGGCTCATGAGACCTTGTGCTGCCCTCAGTGTCT 2618  
 DB 2284 CTGCTTCTCCCTCTGTGTGAGCTCATGAGACCTTCTCTACAGTGAAGCTTACTCT 2343  
 QY 2619 CCAAGTGTGTCAGGCTCATGCTGTGCAAGAGAGGAGCACTTCAACTGACAGACAGA 2678  
 DB 2344 CTAAGTGTGTCAGTGTGCTATTTCAAGGAGACAGAGGAGACATTTGTACTCAACAGCA 2403  
 QY 2679 AAGCCCTGGCATGATGAGGCTGTGTGCGGCGGAGACAGCCCTGCTCCTCGAGACTTC 2738  
 DB 2404 AAGGACTAAGGCTGATGAGGCTGTGTGAGAGAGACACTTCTGCTCTGAGAGACTCC 2463  
 QY 2739 GCTTGGGCTCTCCCATGCTCACTGCTCACTGCTTCCGAGATTAAGTGGGCGCAAGTCCAGA 2798  
 DB 2464 GACTGGGCTCTCTCATGCTCACTGCTTCTCTGGAATTAAGTGAAGAGCAAGCCCAAA 2523  
 QY 2799 AGATGTGGAGCCCATAGCAG 2819  
 DB 2524 AGACCACATGCGCCACAGCAG 2544

RESULT 3  
 BM563641 1084 bp mRNA linear EST 20-FEB-2002  
 LOCUS AGENCOUPT\_6564728 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5744119  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM563641  
 VERSION BM563641.1 GI:18810737  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1084)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.



cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M12764 row: P column: 08  
High quality sequence start: 24  
High quality sequence stop: 713.  
Location/Qualifiers

## FEATURES

source

1. 1084  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5744119"  
/clone\_lib="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: PCMV-SPORT6; Site:1: NotI;  
Site:2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC Library."  
BASE COUNT 328 a 301 c 277 g 176 t 2 others  
ORIGIN

Query Match 24.2%; Score 921; DB 13; Length 1084;  
Best Local Similarity 95.3%; Pred. No. 1.9e-198;  
Matches 1013; Conservative 0; Mismatches 42; Indels 8; Gaps 6;

OY 12 AGAGGTCTAGAGCAATAGCCCTCTCTATTCCTAATCCCTCCCAACCAAGAGGCACA 71  
DB 24 AGAGAGCTAGAGCAATAGCCCTCTCTATTCCTAATCCCTCCCAACCAAGAGGCACA 83  
OY 72 GAGTTCATTACTAGTGGGGCCAGCTGTGATCGGCCAATCGCCAGCTGCTTTAAAGG 131  
DB 84 GAGTTCATTACTAGTGGGGCCAGCTGTGATCGGCCAATCGCCAGCTGCTTTAAAGG 143  
OY 132 AAGACCACTGATGTAGATGATGTAACCAAGAGAGTGCATCATGAGCAATCAA 191  
DB 144 AAGACCACTGATGTAGATGATGTAACCAAGAGAGTGCATCATGAGCAATCAA 203  
OY 192 TGAGAGATCTGAGAGAGAGAGGCTGGTGGAGCCCAAGAGATAGAACTGGAAGT 251  
DB 204 TGAGAGATCTGAGAGAGAGAGGCTGGTGGAGCCCAAGAGATAGAACTGGAAGT 263  
OY 252 CAATATCTCCCTGAGAGGAATTAACATGAGCCAGCTCTAAGTCAAGTCTAGGTCAG 311  
DB 264 CAATATCTCCCTGAGAGGAATTAACATGAGCCAGGTTCTAAGTCAAGTCTAGGTCAG 323  
OY 312 ACTGGCAACCTGAGACACACAGAGGCTATTAACCCCGCTAGAGCTGGGCCAGAAAGA 371  
DB 324 ACTGGCAACCTGAGACACACAGAGGCTATTAACCCCGCTAGAGCTGGGCCAGAAAGA 383  
OY 372 CACCATATGCCAGGACGAGATGAGAGACTCTGCAAGGATCCATACCAAGAGAACCTG 431  
DB 384 CACCATATGCCAGGACGAGATGAGAGACTCTGCAAGGATCCATACCAAGAGAACCTG 443  
OY 432 CTTCGAACCAAGACCCCTACCCAGAGAGACCCCTGCACAACTGATGTGATTCCTCC 491  
DB 444 CTTCGAACCAAGACCCCTACCCAGAGAGACCCCTGCACAACTGATGTGATTCCTCC 503  
OY 492 AGAAGGAACCTAGAGCCCAACAAATCTGCTTCACAGAGAGATTTCTTGGCCACAGA 551  
DB 504 AGAAGGAACCTAGAGCCCAACAAATCTGCTTCACAGAGAGATTTCTTGGCCACAGA 563  
OY 552 AGCCGCAACAGAGATCACTTACATCCAAAGGCTGCTCACTCAAGAGAGAGCG 611  
DB 564 AGCCGCAACAGAGATCACTTACATCCAAAGGCTGCTCACTCAAGAGAGAGCG 623  
OY 612 CCTCCAGCAGGACCTGGGCTAGGAAAAAGAAATCTATTAACAGAGACCAAGCATTTGA 671

|||||  
DB 624 CTTCCCGAGGAGGACCTGGGCTAGAGAAAAGATCTATTAACCAAGAGGACGACATTTGA 683  
OY 672 GACAAAGACATGTAGCCCAAGCCAGGCGCTGGGCCAGAGAGCCACCTCCAGTCAACAG 731  
DB 684 GACAAAGACATGTAGCCCAAGCCAGGCGCTGGGCCAGAGAGCCACCTCCAGTCAACAG 743  
OY 732 AACCTGAATCAACACCTGGGCGCCAGGCTAACTGAGGCCAAAGAGGAGCATCTGCC 791  
DB 744 AACCTGAATCAACACCTGGGCGCCAGGCTAACTGAGGCCAAAGAGGAGCATCTGCC 803  
OY 792 CGACTGATTCATACCTCCCAAGAGACACCTGAAACAGTCAACAAACAAAGCAGATCC 851  
DB 804 CGACTGATTCATACCTCCCAAGAGACACCTGAAACAGTCAACAAACAAAGCAGATCC 863  
OY 852 AGGAGCCCAATCCACAGAGGATCTTTGACAGAGCTGGGATTTCTAACAACCTCAGG 911  
DB 864 AGGAGCCCAATCCACAGAGGATCTTTGACAGAGCTGGGATTTCTAACAACCTCAGG 922  
OY 912 AACATCCATACAGCGATCAG-CCCTAGAGTGAAGGAC-TTTCGATGGGTGCGAGA 969  
DB 923 -ACTATCCCTACAGCGATCAGCCCTTAAGTGGAGGACCTTTCTGATGGGTGCGAGA 981  
OY 970 TTCTGATCAGAAATCAGATGT-GGATCATCTTCAGACAGATTTCTCCAGCCAGATGG 1028  
DB 982 TTCTGATCAGAAATCAGATGTGGGGATCTTCTCCGACCAGATTTCTCCAGCCAGATGG 1041  
OY 1029 GTGGA---TGTTGGCCAGGAGGAGTGAAGCTTCAAGG 1068  
DB 1042 GGGGGAAATGGGGGGCCCAAGGAAATGAAGCTTAAGGCTTCAAGG 1084

RESULT 4  
BI826568 826 bp mRNA linear EST 04-OCT-2001  
LOCUS 603077016F1 NIH\_MGC\_119 Homo sapiens cdna clone IMAGE:516882 5'  
DEFINITION  
mRNA sequence.  
ACCESSION BI826568  
VERSION BI826568.1 GI:15938118  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 826)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@rs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cdna Library Preparation: Life Technologies, Inc.  
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M1419 row: P column: 03  
High quality sequence stop: 823.  
Location/Qualifiers

## FEATURES

source

1. 826  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:516882"  
/clone\_lib="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: PCMV-SPORT6; Site:1: NotI;  
Site:2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH-MGC Library."

BASE COUNT 269 a 231 c 207 g 119 t

Query Match 21.3%; Score 808.6; DB 13; Length 826;  
Best Local Similarity 99.4%; Pred. No. 5.9e-173;  
Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 56 CAAGAAGAGGACGAGTCTACTTACTGATGGGGCCAGCTGTGATCGGCCACTGCC 115  
|||||  
Db 1 CAAGAAGAGGACGAGTCTACTTACTGATGGGGCCAGCTGTGATCGGCCACTGCC 60  
QY 116 AGCTGCTTAAAAAGAGACACAGTATGCTAGATGAGTGAACCCAGAGAAATGC 175  
|||||  
Db 61 AGCTGCTTAAAAAGAGACACAGTATGCTAGATGAGTGAACCCAGAGAAATGC 120  
QY 176 CATCATGAGGAATCAATGAGAGATCTGTGAAGAGAGAGGCTGGTGGAGCCAGAGG 235  
|||||  
Db 121 CATCATGAGGAATCAATGAGAGATCTGTGAAGAGAGAGGCTGGTGGAGCCAGAGG 180  
QY 236 ATGAACTCGGAAGATCAATATCTCCGCTGAGGGAATACAAATGAGCCAGGTTCTAG 295  
|||||  
Db 181 ATGAACTCGGAAGATCAATATCTCCGCTGAGGGAATACAAATGAGCCAGGTTCTAG 240  
QY 296 TCAGTGTAGTCAAGTGGCAACTGAAACACACAGAGGCTTAAACCCGCTAGAG 355  
|||||  
Db 241 TCAGTGTAGTCAAGTGGCAACTGAAACACACAGAGGCTTAAACCCGCTAGAG 300  
QY 356 CCTGGGCCAGAAAAGACACCCATAGCCCAAGCAATCGAAGACTCTGAGGATCCAAAT 415  
|||||  
Db 301 CCTGGGCCAGAAAAGACACCCATAGCCCAAGCAATCGAAGACTCTGAGGATCCAAAT 360  
QY 416 ACCCAACAGAGGCTCTTCAAAACCAAGCCCTCACAGAGAGAGAGCCCTGCACAA 475  
|||||  
Db 361 ACCCAACAGAGGCTCTTCAAAACCAAGCCCTCACAGAGAGAGAGCCCTGCACAA 420  
QY 476 CATGATGCTGAATCCAGAAAGAACTAGAGCCCAACAAAAATCTCTTCAAGAGAA 535  
|||||  
Db 421 CATGATGCTGAATCCAGAAAGAACTAGAGCCCAACAAAAATCTCTTCAAGAGAA 480  
QY 536 TTTTCTGGCCCCACAGAAAGCCGACACAGCAATCACTTACATCCAAAGGCTGCTC 595  
|||||  
Db 481 TTTTCTGGCCCCACAGAAAGCCGACACAGCAATCACTTACATCCAAAGGCTGCTC 540  
QY 596 ACTCAACAGGAAGCTGCTCCAGAGAGGAGCTGGGCTAGGAAAAGATCTATTAACCAA 655  
|||||  
Db 541 ACTCAACAGGAAGCTGCTCCAGAGAGGAGCTGGGCTAGGAAAAGATCTATTAACCAA 600  
QY 656 CAGGAGCCAGCAATTTGAGACAAAGACATGTAGCCAGCCAGAGGCTGGCCAGAGAGCCA 715  
|||||  
Db 601 CAGGAGCCAGCAATTTGAGACAAAGACATGTAGCCAGCCAGAGGCTGGCCAGAGAGCCA 660  
QY 716 CCTCCAGCTCAACAGAAAGTGAATCAACACTGGGCCCCAGGCTTAAACCTGGAGCCAAA 775  
|||||  
Db 661 CCTCCAGCTCAACAGAAAGTGAATCAACACTGGGCCCCAGGCTTAAACCTGGAGCCAAA 720  
QY 776 AGGAGCCATCTGCCCCGAGTGAATCTACATCCCAAGAGACACTTAACAGTGAAGCAAG 835  
|||||  
Db 721 AGGAGCCATCTGCCCCGAGTGAATCTACATCCCAAGAGACACTTAACAGTGAAGCAAG 780  
QY 836 CAAGAAGCCAGTCCAGGAGGAGCCAAATCCAAAGCAGGATCTTTGAC 882  
|||||  
Db 781 CAAGAAGCCAGTCCAGGAGGAGCCAAATCC - AGCAGGACATCTTTGAC 826

RESULT 5  
B1827841 872 bp mRNA linear EST 04-OCT-2001  
LOCUS 603074047f1 NIH\_MGC\_119 Homo sapiens cDNA IMAGE:516131 5',  
DEFINITION mRNA sequence.  
ACCESSION B1827841

VERSION B1827841.1 GI:15939391  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 872)  
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgrpb@remail.nih.gov

Tissue procurement: Life technologies, Inc.  
CDNA library preparation: Life technologies, Inc.  
CDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLNL1412 row: m column: 12  
High quality sequence stop: 869.

FEATURES  
source Location/Qualifiers  
1..872  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:516131"  
/clone\_1lb="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"

/note="Organ: Drain; Vector: pcMV-SpOrf6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb. Insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH-MGC Library."  
BASE COUNT 279 a 244 c 217 g 132 t

Query Match 20.7%; Score 787; DB 13; Length 872;  
Best Local Similarity 97.6%; Pred. No. 4.9e-168;  
Matches 852; Conservative 0; Mismatches 15; Indels 6; Gaps 5;

QY 66 GGCACAGATGCTACTTACTAGTGGGGCCAGCTGTGATGGCCAACTGCGACCTGCTTA 125  
|||||  
Db 1 GGCACAGATGCTACTTACTAGTGGGGCCAGCTGTGATGGCCAACTGCGACCTGCTTA 60  
QY 126 AAAAGGAGACCAAGTATCTAGGATGAGTGAAGAAACCAAGAGAGATGCCATGAGG 185  
|||||  
Db 61 GAAAGGAGACCAAGTATCTAGGATGAGTGAAGAAACCAAGAGAGATGCCATGAGG 120  
QY 186 AATCAATGAGATCTGTGAAGAGAGAGGCTGGGCGGAGCCCAAGAGATGAACCTG 245  
|||||  
Db 121 AATCAATGAGATCTGTGAAGAGAGAGGCTGGGCGGAGCCCAAGAGATGAACCTG 180  
QY 246 GAAAGTCAATATCTCCCGTGAAGGAAATTAACATGAGAGCCAGTCTTAAGTGAAGTCTTA 305  
|||||  
Db 181 GAAAGTCAATATCTCCCGTGAAGGAAATTAACATGAGAGCCAGTCTTAAGTGAAGTCTTA 240  
QY 306 GGTCAAGCTGGCAACCTGACACACACAGAGGCTTAAACCCGCTAGAGCTGGGCCAG 365  
|||||  
Db 241 GGTCAAGCTGGCAACCTGACACACACAGAGGCTTAAACCCGCTAGAGCTGGGCCAG 300  
QY 366 AAAAGACACCCATAGCCCAAGAGATGGAAGACTCTGAGAGATCCATACCAACAGAG 425  
|||||  
Db 301 AAAAGACACCCATAGCCCAAGAGATGGAAGACTCTGAGAGATCCATACCAACAGAG 360  
QY 426 AGCTGCTTCAAAACCAAGAGAGCCCTCACAGAGAGAGAGCCCTGCACACATGATGCTG 485  
|||||  
Db 361 AGCTGCTTCAAAACCAAGAGAGCCCTCACAGAGAGAGAGCCCTGCACACATGATGCTG 420

486 AATCCAGAGGAAAGCTAGAGCCCAACAAAATCTGCTTCACAGAGATTTCTGCC 545  
 Db 421 AATCCAGAGGAAAGCTAGAGCCCAACAAAATCTGCTTCACAGAGATTTCTGCC 480  
 546 CACAGAGGCGCCGACACAGCAATCACTTACATCCAAAGGGTGTCTCCTCAGAG 605  
 Db 481 CACAGAGGCGCCGACACAGCAATCACTTACATCCAAAGGGTGTCTCCTCAGAG 540  
 606 AAGCTCCCTCCAGAGGAGGAGCTGGGCTAGAGAAAAGATTTATTAATCAAGAGCCAG 665  
 Db 541 AAGCTCCCTCCAGAGGAGGAGCTGGGCTAGAGAAAAGATTTATTAATCAAGAGCCAG 600  
 666 CATTGAGACAAAGACATGTGA-GGCCAGCCAGGGCTGGGCGAGAGAGCCACCTCCAGCT 724  
 Db 601 CATTGAGACAAAGACATGTGA-GGCCAGCCAGGGCTGGGCGAGAGAGCCACCTCCAGCT 660  
 725 CAACAGAAAGCTGAATCAACACCTGGGCGAGGCTTAACCTGAGCCAAAGGAGGAGCA 784  
 Db 661 CAACAGAAAGCTGAATCAACACCTGGGCGAGGCTTAACCTGAGCCAAAGGAGGAGCA 719  
 785 TCTG-CCCCGAGTGAATCTTACATCCCAAGAGACCTGAGACATCAAGCAAGCAAC 843  
 Db 720 TCTG-CCCCGAGTGAATCTTACATCCCAAGAGACCTGAGACATCAAGCAAGCAAC 779  
 844 GCCAGTCCAGAGGAGCCAAATCCAGAGGAGG--ATCTTGAAGAGCTGGGATTTCTAACA 901  
 Db 780 GGGAGTCCAGAGGAGCCAAATCCAGAGGAGGATTTCTTGAAGAGAGCTGGGATTTCTAC- 838  
 902 AACTTTCAGAACTATTCATACAGGATCAGCC 934  
 Db 839 AACTTTCAGAACTATTCATACAGGATCAGCC 871

RESULT 6  
 BI827706 905 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603074127F1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:516595 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI827706  
 VERSION BI827706.1 GI:15939256  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 905)  
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bhs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.llnl.gov  
 Plate: LAM1412 row: f column: 08  
 High quality sequence stop: 890.  
 Location/Qualifiers  
 1..905  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:516595"  
 /clone\_1db="NIH\_MGC\_119"  
 /tissue\_type="medulla"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-Sport6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source normal medulla from  
 anonymous male age 27. Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.3 kb, insert size range

0.9-3 kb. Library is normalized and enriched for  
 (full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 013. Note:  
 this is a NIH-MGC Library."

BASE COUNT 291 a 255 c 216 g 143 t  
 ORIGIN

Query Match 20.6%; Score 784; DB 13; Length 905;  
 Best Local Similarity 96.7%; Pred. No. 2,4e-167;  
 Matches 876; Conservative 0; Mismatches 20; Indels 10; Gaps 7;

21 GGCACATAGCCCTCTCATTTCTAATCTCCACCAAGAGAGGACAGTTCATT 80  
 Db 1 GGCACATAGCCCTCTCATTTCTAATCTCCACCAAGAGAGGACAGTTCATT 60  
 81 ACTTAGTGGGGGACGCTGTATGCGCCACTGCCAGTGCCTTAAAAAGAGACAGT 140  
 Db 61 ACTTAGTGGGGGACGCTGTATGCGCCACTGCCAGTGCCTTAAAAAGAGACAGT 120  
 141 GATGCTAGGATGAGTGAACCCAAAGAGAGTGCATCATGAGGAATCATGAGATC 200  
 Db 121 GATGCTAGGATGAGTGAACCCAAAGAGAGTGCATCATGAGGAATCATGAGATC 180  
 201 TGTGAAGAGAGGGGCTGGGTGGAGAGCCAGAGGATGAACCTGGAAGATCATATCTC 260  
 Db 181 TGTGAAGAGAGGGGCTGGGTGGAGAGCCAGAGGATGAACCTGGAAGATCATATCTC 240  
 261 CCGTAGGGAATTAACATGAGGAGCTTTAAGTCTAGTCTAGTCTAGTCTAGTCT 320  
 Db 241 CCGTAGGGAATTAACATGAGGAGCTTTAAGTCTAGTCTAGTCTAGTCTAGTCT 300  
 321 CTGAACCAACACAGAGGCTTAACCCGCTAGAGGCTGGGCGAAGAACACCCATAG 380  
 Db 301 CTGAACCAACACAGAGGCTTAACCCGCTAGAGGCTGGGCGAAGAACACCCATAG 360  
 381 CCCAGCAGATTCAGAGAGCTGAGAGGATCAATATCCCAAGAGGCTCTTCAACC 440  
 Db 361 CCCAGCAGATTCAGAGAGCTGAGAGGATCAATATCCCAAGAGGCTCTTCAACC 420  
 441 AAGAGCCCTTACCCAGAGAGAGCCCTGCACAAATATGCTGAATCCAGAGAGAG 500  
 Db 421 AAGAGCCCTTACCCAGAGAGAGCCCTGCACAAATATGCTGAATCCAGAGAGAG 480  
 501 CTAGAGCCCAACAAATCTCTTCTACAGAGGATTTCTGGCCCAAGAGAGCCGAG 560  
 Db 481 CTAGAGCCCAACAAATCTCTTCTACAGAGGATTTCTGGCCCAAGAGAGCCGAG 540  
 561 CACAGCAATCACTTACATCCAAAGGGTGTCTCACTCAACAGAGAGGCTCTCCAGC 620  
 Db 541 CACAGCAATCACTTACATCCAAAGGGTGTCTCACTCAACAGAGAGGCTCTCCAGC 600  
 621 AGGAGCTGGGCTAGAGAAAGATCTATTAATCAACAGAGGCTCTCCAGC 680  
 Db 601 AGGAGCTGGGCTAGAGAAAGATCTATTAATCAACAGAGGCTCTCCAGC 660  
 681 ATGAGCCCAACAGAGGCTGGGCTGGAGAGGAGCCAGCTCAACAAAGAGTGAAT 740  
 Db 661 ATGAGCCCAACAGAGGCTGGGCTGGAGAGGAGCCAGCTCAACAAAGAGTGAAT 720  
 741 CAACAGCTGGGCGGAGGCTTAACCTGAGCCAAAGAGGAGCC--ATCTGCCGAGCTG 797  
 Db 719 CAACAGCTGGGCGGAGGCTTAACCTGAGCCAAAGAGGAGGAGCCATCTGCTCCGAGCTTG 778  
 798 AATCTACATCCAA-GAGACAGCTGAA--CAGTCAGACAGCAACAAAGCAGCTC-AG 853  
 Db 779 AATCTACATCCAAAGACAGCTGAA--CAGTCAGACAGCAACAAAGCAGCTC-AG 838  
 854 GGAGCCAAATCCAGAGGAGTCTTGAACAGAGCTGGGATTTCTAACA-AACTTCAGGA 912  
 Db 839 GGAGCCAAATCCAGAGGAGTCTTGAACAGAGCTGGGATTTCTAACA-AACTTCAGGA 898  
 913 ACTATC 918  
 |||||

Db 899 ACTATC 904

RESULT 7  
BI827559  
LOCUS  
DEFINITION BI827559 870 bp mRNA linear EST 04-OCT-2001  
603073524F1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5165788 5',  
mRNA sequence.  
ACCESSION BI827559  
VERSION BI827559.1 GI:15939096  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 870)  
NIH-MGC <http://mgc.nci.nih.gov/>,  
National Institutes of Health, Mammalian Gene Collection (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1141 row: 0 column: 05  
High quality sequence stop: 848.  
FEATURES  
source  
1. 870  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5165788"  
/clone\_lib="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1; Note:  
Site 2: Ecorev (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (Ecorev site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC library."

BASE COUNT 287 a 234 c 221 g 128 t  
ORIGIN

Query Match 20.6%; Score 782.6; DB 13; Length 870;  
Best Local Similarity 97.8%; Pred. No. 4.9e-167;  
Matches 846; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

QY 55 CCAAGAAAGAGCAGCAGATTCATTAGTGGGGGCCAGCTGTATGGCCACT-G 113  
DB 1 CCAAGAAAGAGCAGCAGATTCATTAGTGGGGGCCAGCTGTATGGCCACTGG 60

QY 114 CCAAGCTGCCCTTAAAGAGACAGCAGATGCTAGATGAGTGAACCCAGAGAGT 173  
DB 61 GGGGCTGCCCTTAAAGAGACAGCAGATGCTAGATGAGTGAACCCAGAGAGT 120

QY 174 GCCATCATGAGATCAATGAGATCTGTGAAGAGAGAGGGCTGGTGGAGCCAGAA 233  
DB 121 GCCATCATGAGATCAATGAGATCTGTGAAGAGAGAGGGCTGGTGGAGCCAGAA 180

QY 234 GGATGAGACCTGGAATCAATATCTCCGTGAGGGAATPAACATGAGCCAGTTCTA 293  
DB 181 GGATGAGACCTGGAATCAATATCTCCGTGAGGGAATPAACATGAGCCAGTTCTA 240

QY 294 AGTCAGTGTAGTGTAGTGTGCAACCTGACACACACAGAGCCCTATTAACCCGCTAG 353  
DB 241 AGTCAGTGTAGTGTAGTGTGCAACCTGACACACAGAGCCCTATTAACCCGCTAG 300

QY 354 AGCCTGGCCAGAAAGACACCCATAGCCAGCCAGAACTGTCAGGATCCA 413  
DB 301 AGCCTGGCCAGAAAGACACCCATAGCCAGCCAGAACTGTCAGGATCCA 360

QY 414 ATACCAAGAGAAAGCTGCTTCAAAACCAAGACCCCTACCCAGAGAGACCCCTGCAC 473  
DB 361 ATACCAAGAGAAAGCTGCTTCAAAACCAAGACCCCTACCCAGAGAGACCCCTGCAC 420

QY 474 AACATGATGCTGAATCCCAAGAGAACCTTAGAGCCCAAGAAATGCTTCAAGAGG 533  
DB 421 AACATGATGCTGAATCCCAAGAGAACCTTAGAGCCCAAGAAATGCTTCAAGAGG 480

QY 534 AATTTCTTGGCCCAAGAGACCCGACACAGCAATCCTTACATCCAAAGGCTGTC 593  
DB 481 AATTTCTTGGCCCAAGAGACCCGACACAGCAATCCTTACATCCAAAGGCTGTC 540

QY 594 TCACTCAACAGAAAGCTGCTCCAGCAGAGGACCTGGGCTAGAGAAAGATCTATAAC-T 652  
DB 541 TCACTCAACAGAAAGCTGCTCCAGCAGAGGACCTGGGCTAGAGAAAGATCTATAACGT 600

QY 653 CAACAGAGGACAGATTCAGCAAAAGATGTAGCCAGCCAGGACCTGGGCTAGAGAG 712  
DB 601 CAACAGAGGACAGATTCAGCAAAAGATGTAGCCAGCCAGGACCTGGGCTAGAGAG 660

QY 713 CCAACTCCAGCTCAACAGAAAGCTGATCAA-CACCTGCGCCAGGCTAAACCTGGAGC 771  
DB 661 CCAACTCCAGCTCAACAGAAAGCTGATCAAAGCAGCTGGGCTAAACCTGGAGC 720

QY 772 CAAAAGGAGCCATCTGCCCCGACTGATATCAATCCCAAGACACCTTAACAGTCAGA 831  
DB 721 CAAAAGGAGCCATCTG-CGCCAGCTGATCTAGCT-CCAAGACACCTTAACAGTCAGA 778

QY 832 CAAGCAAAAGAGCAGATTCAGGAGAGCCAAATCAAGCAGAGATCTTGACAGAGCTGGG 891  
DB 779 CAAGCAAAAGAGCAGATTCAGGAGAGCCAAATCAAGCAGAGATCTTGACAGAGCTGGG 838

QY 892 ATTTCTAACAATACTTCAGAACTA 916  
DB 839 ATTTCTAACAATACTTCAGAACTA 863

RESULT 8  
BI561175 871 bp mRNA linear EST 05-SEP-2001  
LOCUS BI561175  
DEFINITION 603253712F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5296155 5',  
mRNA sequence.  
ACCESSION BI561175  
VERSION BI561175.1 GI:15448489  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 871)  
NIH-MGC <http://mgc.nci.nih.gov/>,  
National Institutes of Health, Mammalian Gene Collection (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11749 row: 9 column: 04  
High quality sequence stop: 820.  
FEATURES  
source  
1. 871  
Location/Qualifiers  
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:5296155"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/Note="Organ: testis; Vector: BluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carlncl, in preparation). Library
constructed by M. Brownstein (NIMH/NHGR), National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      281 a      242 c      215 g      133 t
ORIGIN

```

```

Query Match      20.1%; Score 764.8; DB 13; Length 871;
Best Local Similarity 96.6%; Pred. No. 5.5e-163;
Matches 835; Conservative 0; Mismatches 22; Indels 7; Gaps 5;

OY 1 CTTCTGTAGAGAGTGTGAGGACATAGCCCTCTCTATCTCTATCTCTCCACCAAG 60
Db 6 CTTCTGTAGAGAGTGTGAGGACATAGCCCTCTCTATCTCTATCTCTCCACCAAG 65
OY 61 AAGAGGACACAGTTCATTTACTAGTGGGGCCACGCTGTGATCGCCAACTGCCAGT 120
Db 66 AAGAGGACACAGTTCATTTACTAGTGGGGCCACGCTGTGATCGCCAACTGCCAGT 125
OY 121 CCTTAAAAAGAGACCACTGATGCTAGATGAGTGAACCCAGAGAGAGTCCATCA 180
Db 126 CCTTAAAAAGAGACCACTGATGCTAGATGAGTGAACCCAGAGAGAGTCCATCA 185
OY 181 TGAGGAATCAATGAGATCTGTGAAGAGAGGCTGTGGTGGAGCCAGAGAGTATGA 240
Db 186 TGAGGAATCAATGAGATCTGTGAAGAGAGGCTGTGGTGGAGCCAGAGAGTATGA 245
OY 241 ACCTGGAAGATCAATATCTCCGTGAGGAGAAATCAATGAGAGCCAGTTCATGTCAGT 300
Db 246 ACCTGGAAGATCAATATCTCCGTGAGGAGAAATCAATGAGAGCCAGTTCATGTCAGT 305
OY 301 GCTTGGGTGAGACTGGCAACCTGAAACACAGAGGCTGTATTAACCCGCTAGAGCTGG 360
Db 306 GCTTGGGTGAGACTGGCAACCTGAAACACAGAGGCTGTATTAACCCGCTAGAGCTGG 365
OY 361 GCCAAGAAAAGAACCCATGAGCCGACGAGAAATGAGAGACTGTGACAGAGTCCATATCA 420
Db 366 GCCAAGAAAAGAACCCATGAGCCGACGAGAAATGAGAGACTGTGACAGAGTCCATATCA 425
OY 421 ACAGAGAGCTGCTTCAAAACCAAGACCCCTCAACGAGAGAGCCCTGACACATCATGA 480
Db 426 ACAGAGAGCTGCTTCAAAACCAAGACCCCTCAACGAGAGAGCCCTGACACATCATGA 485
OY 481 TGCTGAATCCAGAAAGAACTAGAGCCCAACAAAATCTGCTTCAAGAGAGATTTCT 540
Db 486 TGCTGAATCCAGAAAGAACTAGAGCCCAACAAAATCTGCTTCAAGAGAGATTTCT 545
OY 541 TGCCCCAAGAGAGCCGACACACCAATCACTTATCAATCCAAAGGCTGC-TGCTCACTC 599
Db 546 TGCCCCAAGAGAGCCGACACACCAATCACTTATCAATCCAAAGGCTGCTTCACTC 605
OY 600 AACAGAGAGCTGCTTCCAGAGAGGAGCTGGGCTAGAGAAAGATCTATATCAACACAG 659
Db 606 AACAGAGAGCTGCTTCCAGAGAGGAGCTGGGCTAGAGAAAGATCTATATCAACACAG 665
OY 660 AGCCAGCATTTGAGCAAAAGACATGTAGCCAGGCCGCTGGGGCCAGAGGAGAGCCA 725
Db 666 AGCCAGCATTTGAGCAAAAGACATGTAGCCAGGCCGCTGGGGCCAGAGGAGAGCCA 725
OY 717 CTCGAGCTCAACAAGAGTGAATCAACACCTGCGGCCAGAGCTTAACCTGGAGCCAAAA 776
Db 726 CTCGAGCTCAACAAGAGTGAATCAACACCTGCGGCCAGAGCTTAACCTGGAGGCC-AAA 784
OY 777 GGGAGCATCTTCCCGACTGAATCTACATCCAGAGAGACACTGTAACGTGAGACAAG 836

```

```

Db 785 GGGAGCATCTTCCCGACTG-ATCTAGCTGCCAAGAGA-ACCTGACAGTCACAGCAGC 842
OY 837 AACACAGCCAGTCCAGGAGACCA 860
Db 843 AACACAGCCAGTCCAGGAGACCA 866

RESULT 9
B1830913
LOCUS
DEFINITION
B1830913 787 bp mRNA linear EST 04-OCT-2001
603081040F1 NIH_MGC_119 Homo sapiens CDNA clone IMAGE:5172728 5',
mRNA sequence.
B1830913
VERSION
B1830913.1 GI:15942463
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 787)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgrabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: L14M1429 row: P column: 09
High quality sequence stop: 784.
Location/Qualifiers
1..787
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5172728"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT      254 a      219 c      200 g      114 t
ORIGIN
Query Match      19.9%; Score 757; DB 13; Length 787;
Best Local Similarity 99.1%; Pred. No. 3.1e-161;
Matches 782; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

OY 66 GGCACAGAGTTCATTTACTTGTGGGGCCAGCTGTGATCGGCACTGCGCTTGA 125
Db 1 GGCACAGAGTTCATTTACTTGTGGGGCCAGCTGTGATCGGCACTGCGCTTGA 60
OY 126 AAAAGGAAGACAGTGTAGTGAAGTGAACCCAGAGAGAGTGCATCATGAGG 185
Db 61 GAAAGGAAGACAGTGTAGTGAAGTGAACCCAGAGAGAGTGCATCATGAGG 120
OY 186 AATCAATGAGAGATCTGTGAAGAGAGAGGCTGGGTGGAGCCAGAGAGATGACCTG 245
Db 121 AATCAATGAGAGATCTGTGAAGAGAGAGGCTGGGTGGAGCCAGAGAGATGACCTG 180
OY 246 GAAGATCAATATCTCCGCTAGAGGAAATTAACATGAGAGGCTTCAATCACTGCTCA 305

```

```

Db      181 GAAGATCAATATCTCCCGTGAGGGAATTAACAAATGAGCCAGGTTCTAGTCAATGTCTA 240
QY      306 GGTGAGCTGTCGAACCTGGAACCAACACAGAGGCTTAAACCCGCTTAGAGCTGGGCCAG 365
Db      241 GGTGAGCTGTCGAACCTGGAACCAACACAGAGGCTTAAACCCGCTTAGAGCTGGGCCAG 300
QY      366 AAAAGACACCCATAGGCCACCCAGAAATGCAAGACTGTCAGGATCCAAATACCAACAG 425
Db      301 AAAAGACACCCATAGGCCACCCAGAAATGCAAGACTGTCAGGATCCAAATACCAACAG 360
QY      426 AGCTGCTTCAACCAAAAGACCCCTCAACCCAGAGAGACCCCTGCACAAATGATGCTG 485
Db      361 AGCTGCTTCAACCAAAAGACCCCTCAACCCAGAGAGACCCCTGCACAAATGATGCTG 420
QY      486 AATCCGAGAGGAAGCACTAGAGCCCAACAAATCTGCTTCAACAGAGGAATTTCTGCCC 545
Db      421 AATCCGAGAGGAAGCACTAGAGCCCAACAAATCTGCTTCAACAGAGGAATTTCTGCCC 480
QY      546 CACAGAGACCCGACACCAAGCAATCACTTACATCCAAAGGGTGTCTCACTCAACAG 605
Db      481 CACAGAGACCCGACACCAAGCAATCACTTACATCCAAAGGGTGTCTCACTCAACAG 540
QY      606 AAGCTGCTCCGAGAGGAGCACTGGGCTTGAAGAAATCTTAATCAACAGAGCCAG 665
Db      541 AAGCTG-CTCCGAGAGGAGCACTGGGCTTGAAGAAATCTTAATCAACAGAGCCAG 599
QY      666 CATTGAGACAAAGACATGATGACCCAGAGGCTTGGGCCAGAGAGCACTCCAGCTC 725
Db      600 CATTGAGACAAAGACATGATGACCCAGAGGCTTGGGCCAGAGAGCACTCCAGCTC 658
QY      726 AACAGAAAGCTGAATCAACACCTGGGGCCAGAGCTTAACCTTGGAGCCAAAGGGAGCAT 785
Db      659 AACAGAAAGCTGAATCAACACCTGGGGCCAGAGCTTAACCTTGGAGCCAAAGGGAGCAT 718
QY      786 CTGCCCCGACTGAATCTACATCCCAAGAGACACCTGAACAGTCAGACAGCAACAAACAGC 845
Db      719 CTGCCCCGACTGAATCTACATCCCAAGAGACACCTGAACAGTCAGACAGCAACAAACAGC 778
QY      846 CAGTCCAGG 854
Db      779 CAGTCCAGG 787

RESULT 10
LOCUS   B1520246 788 bp mRNA linear EST 29-AUG-2001
DEFINITION 603071104.F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163288 5',
mRNA sequence.
ACCESSION B1520246
VERSION   B1520246.1 GI:15345038
KEYWORDS EST.
SOURCE    human.
ORGANISM Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 788)
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished (1999)
COMMENT  Contact: Robert Strausberg, Ph.D.
            Email: c9apds-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM1405 row: 9 column: 01
            High quality sequence stop: 775.
            Location/Qualifiers
                1..788
                /organism="Homo sapiens"

```

```

Query Match 19.8% Score 753.2; DB 13; Length 788;
Best Local Similarity 99.4%; Pred. No. 2.3e-160;
Matches 777; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Db      1046 GGAGTGAAGCTAGAGCTTCAAGAAATCTGTTATAAAGTGTGATGATACAGTGGG 1105
QY      3 GGAGTGAAGCTAGAGCTTCAAGAAATCTGTTATAAAGTGTGATGATACAGTGGG 62
Db      1106 ACGTCCGACATGAGAAAACCAAGTGTGGAATACAGACACTACCAAGATACAGCTCA 1165
QY      63 ACGTCCGACATGAGAAAACCAAGTGTGGAATACAGACACTACCAAGATACAGCTCA 122
Db      1166 A-GGCTCATCCACAACATGAGACCTGCGCACAAATGACACAGTGTGCTGGTGTGAGGCGGA 1224
QY      123 AGGGCTCATCCACAACATGAGACCTGCGCACAAATGACACAGTGTGCTGGTGTGAGGCGGA 182
Db      1225 GGACAACATACCTTCTTCTGAGCCAGAGGCTCTGGGGAAAACGGCCAGCGGCTGTCAGG 1284
QY      183 GGACAACATACCTTCTTCTGAGCCAGAGGCTCTGGGGAAAACGGCCAGCGGCTGTCAGG 241
Db      1285 CGTTTTCGCGGTAGCGGAGAGCGGCTGGGGCTGGAGCCGGCCCTGGGGCCGCTGCT 1344
QY      242 CGTTTTCGCGGTAGCGGAGAGCGGCTGGGGCTGGAGCCGGCCCTGGGGCCGCTGCT 301
Db      1345 GGGTGTGGCGACCTCTTTGACCTGAGACCAAGACACCGGCAAGCGGATACCGGACCT 1404
QY      302 GGGTGTGGCGACCTCTTTGACCTGAGACCAAGACACCGGCAAGCGGATACCGGACCT 361
Db      1405 AGTGCACACAGCCCGCGTCTGGGCGACCTCTGACAAATCCCGCTATGTGGCTTC 1464
QY      362 AGTGCACACAGCCCGCGTCTGGGCGACCTCTGACAAATCCCGCTATGTGGCTTC 421
Db      1465 CAACCGCGGAGCATCTTCTTCGCAACAGCAACACTGGCGAGCTGAGAGCTTACCT 1524
QY      422 CAACCGCGGAGCATCTTCTTCGCAACAGCAACACTGGCGAGCTGAGAGCTTACCT 481
Db      1525 GGGTGGCCGTACACCACTGCGGCGCTGGTCTACTACGCCAGCGCTGCTGTTACCAA 1584
QY      482 GGGTGGCCGTACACCACTGCGGCGCTGGTCTACTACGCCAGCGCTGCTGTTACCAA 541
Db      1585 TCGGCGGGGGGTACTCTTCTTGAAGGCGCAGAGGGGCTCACGGCGAGCTTCTCGGGA 1644
QY      542 TCGGCGGGGGGTACTCTTCTTGAAGGCGCAGAGGGGCTCACGGCGAGCTTCTCGGGA 601
Db      1645 GTATGTACAGCTGACATTAAGGATGCTTCTATGCGCGCTGCTGCTTCACTTACGCG 1704
QY      602 GTATGTACAGCTGACATTAAGGATGCTTCTATGCGCGCTGCTGCTTCACTTACGCG 661
Db      1705 TGGCATCGGCGCATCTCTCAACACATCTTCATTGGGCTGTGTCTTCGGGAGCACTA 1764
QY      662 TGGCATCGGCGCATCTCTCAACACATCTTCATTGGGCTGTGTCTTCGGGAGCACTA 721
Db      1765 CAAAGCAACGAGACAGGCTCAGTGTGGCGCAGCTCTCTTACACAGGCGCGCTT 1824
QY      722 CAAAGCAACGAGACAGGCTCAGTGTGGCGCAGCTCTCTTCTTACACAGGCGCGCT 781

```

```

/db_xref="taxon:9606"
/clone="IMAGE:5163288"
/clone.lib="NIH_MGC_119"
/tissue.type="medulla"
/lab.host="DH10B"
/Note="Organ: brain; Vector: pCMV-Sport6; site_1: NotI;
site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT 154 a 258 c 226 g 150 t
ORIGIN

```

OY 1825 TG 1826  
Db 782 AG 783

RESULT 11  
B1520160 773 bp mRNA linear EST 29-AUG-2001  
LOCUS 603071405F1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5163360 5',  
DEFINITION mRNA sequence.  
ACCESSION B1520160  
VERSION B1520160.1 GI:15344952  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 773)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM11405 row: j column: 01  
High quality sequence stop: 769.

FEATURES  
source  
1. 773  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5163360"  
/clone\_1lb="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb. Insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH-MGC Library."

BASE COUNT 152 a 254 c 220 g 147 t

ORIGIN  
Query Match 19.1%; Score 726; DB 13; Length 773;  
Best Local Similarity 99.5%; Pred. No. 3.4e-154;  
Matches 770; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

OY 1043 CAGGAGTGAAGCTAGCTCAAGGAATCTGTTATAGTGTGACAGTACAGT 1102  
Db 1 CAGGAGTGAAGCTAGCTCAAGGAATCTGTTAT-AGGTGATGTGAGATACAGT 59  
OY 1103 GAGAGTGGCCACATGAGAAACAGTGTCTGGAATCACAGACATACAGAGTACAGCC 1162  
Db 60 GGGAGTGGCCACATGAGAAACAGTGTCTGGAATCACAGACATACAGAGTACAGCC 119  
OY 1163 TCAAGCTTCATCCACACATGAGACTGCGCACATGACACAGTGTCTGTGACTGTGGCG 1222  
Db 120 TCAAGCTTCATCCACACATGAGACTGCGCACATGACACAGTGTCTGTGACTGTGGCG 179  
OY 1223 GAGGCAACATAGCTCTTCTCCGAGCCAGAGTGTCTGGGAAACGGCCAGGGCTGTCA 1282  
Db 180 GAGGCAACATAGCTCTTCTCCGAGCC-GGTCTGGGGGAAACGGCCAGGGCTGTCA 238  
OY 1283 GGCGTTTTGGCGGTGTACGGGAGCAGGGCGGTGGGGCTGGAGCGGCGCTGTGGCGCTG 1342

Db 239 GCGGTTTTGCCGGTGTACGGGAGCAGCGCTGGGGCTGGAGCCGGCCCTGGCGGCTG 298  
OY 1343 CTGGGTGTGGCGCACCTCTTTGACCTTGACCTGACACCCAGACACCGGCCAACCGGTACCAGAC 1402  
Db 299 CTGGGTGTGGCGCACCTCTTTGACCTTGACCTGACACCCAGACACCGGCCAACCGGTACCAGAC 358  
OY 1403 CTAGTGCACACAGCCCGCTGCTGGCGACACTCTGACCAATATCCCGTATGTGGCC 1462  
Db 359 CTAGTGCACACAGCCCGCTGCTGGCGACACTCTGACCAATATCCCGTATGTGGCC 418  
OY 1463 TCCAAACCGCCGACAGCATCTTCTCCGACACAGCCACCAACCTGGCCGAGCTGAGAGCTTAC 1522  
Db 419 TCCAAACCGCCGACAGCATCTTCTCCGACACAGCCACCAACCTGGCCGAGCTGAGAGCTTAC 478  
OY 1523 CTGGCTCCCTCACCCAGCTCCCGGCTGTGTCTACTACGCCGCCGCTGTGTTAAC 1582  
Db 479 CTGGCTCCCTCACCCAGCTCCCGGCTGTGTCTACTACGCCGCCGCTGTGTTAAC 538  
OY 1583 AATCGGCGGGGCTACTCTTGTAGGGGAGAGAGGGGCTCACCGCGACTTCTCCGG 1642  
Db 539 AATCGGCGGGGCTACTCTTGTAGGGGAGAGAGGGGCTCACCGCGACTTCTCCGG 598  
OY 1643 GAGTA-TGTACAGCTGATTAAGGATCTTATAGCGCTGCTGGGCTTCCAGTTAC 1701  
Db 599 GAGTATGTGACCGCTGATTAAGGATCTTATAGCGCGCTGCTGGGCTTCCAGTTAC 658  
OY 1702 GCCTGCCATCCGGCATCTCTGACAGACCATTCCTATTTGGGCTGTGTGGGGAGCA 1761  
Db 659 GCCTGCCATCCGGCATCTCTGACAGACCATTCATTTGGGCTGTGTGGGGAGCA 718  
OY 1762 CTCAAAAGCAAGAGAG-ACAGGCTCAGTGTGGCGGCGAGCTCTCTTACCA 1814  
Db 719 CTCAAAAGCAAGAGAGAGAGGCTCAGTGTGGCGGCGAGCTCTCTTACCA 772

RESULT 12  
B1832284 781 bp mRNA linear EST 04-OCT-2001  
LOCUS B1832284  
DEFINITION 603080125F1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5171623 5',  
LOCUS B1832284  
DEFINITION mRNA sequence.  
ACCESSION B1832284  
VERSION B1832284.1 GI:15943834  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 781)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapds-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM11427 row: b column: 08  
High quality sequence stop: 781.

FEATURES  
source  
1. 781  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5171623"  
/clone\_1lb="NIH\_MGC\_119"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source normal medulla from



anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics Tracking code 013. Note: this is a NIH-MGC Library."

BASE COUNT 250 a 216 c 195 g 120 t  
ORIGIN

Query Match 18.7%; Score 711.2; DB 13; Length 781;  
Best Local Similarity 97.2%; Pred. No. 7.9e-151;  
Matches 756; Conservative 0; Mismatches 18; Indels 4; Gaps 3;

15 GTCTAGGACATAGCCCTCTCTATTCCTATCTCCACCAAG-AGAAAGAGCAGACA 73  
|||||  
Db GTCTAGGACATAGCCCTCTCTATTCCTATCTCCACCAAGAGAGCAGACA 65  
74 GTTCATTACTTACTTACTGGGGCCAGCTGTGATGGCCAACTGCCAGTCCCTTAAAGAA 133  
|||||  
Db GTTCATTACTTACTTACTGGGGCCAGCTGTGATGGCCAACTGCCAGTCCCTTAAAGAA 125  
134 GACCAATGATGCTAGATGAGTGAAGCCCAAGAGAGGATCATGATGAGATCATG 193  
|||||  
Db GACCAATGATGCTAGATGAGTGAAGCCCAAGAGAGGATCATGATGAGATCATG 185  
126 GACCAATGATGCTAGATGAGTGAAGCCCAAGAGAGGATCATGATGAGATCATG 195  
194 AGAGATCTGTGAAGAGAGAGGCTGGGTGGAGCCCAAGAGATAGAACTGGAGATCA 253  
|||||  
Db AGAGATCTGTGAAGAGAGAGGCTGGGTGGAGCCCAAGAGATAGAACTGGAGATCA 245  
186 AGAGATCTGTGAAGAGAGAGGCTGGGTGGAGCCCAAGAGATAGAACTGGAGATCA 245  
254 ATATCTCCCTGTAGGAGAAATACAAATGAGCCAGGTTCTTAAGTCACTGTAGTCA 313  
|||||  
Db ATATCTCCCTGTAGGAGAAATACAAATGAGCCAGGTTCTTAAGTCACTGTAGTCA 305  
246 ATATCTCCCTGTAGGAGAAATACAAATGAGCCAGGTTCTTAAGTCACTGTAGTCA 305  
314 TGGCACTGTAACACACAGAGAGGCTATTAACCCGCTAGAGCTGGGCGAGAAAGACA 373  
|||||  
Db TGGCACTGTAACACACAGAGAGGCTATTAACCCGCTAGAGCTGGGCGAGAAAGACA 365  
306 TGGCACTGTAACACACAGAGAGGCTATTAACCCGCTAGAGCTGGGCGAGAAAGACA 365  
374 CCCATAGCCAGCCAGCAATGCAAGACTGTGACAGGATTCATACCAACAGAGAGCTGT 433  
|||||  
Db CCCATAGCCAGCCAGCAATGCAAGACTGTGACAGGATTCATACCAACAGAGAGCTGT 425  
366 CCCATAGCCAGCCAGCAATGCAAGACTGTGACAGGATTCATACCAACAGAGAGCTGT 425  
434 TCAACCAAGAGCCCTCAACCCAGAGAGAGCCCTGACACATGATGCTGATCCAG 493  
|||||  
Db TCAACCAAGAGCCCTCAACCCAGAGAGAGCCCTGACACATGATGCTGATCCAG 485  
426 TCAACCAAGAGCCCTCAACCCAGAGAGAGCCCTGACACATGATGCTGATCCAG 485  
494 AAGGATCTGAGAGCCCAAAATCTGCTCAAGAGAGAAATTTCTTCCCAACAGAG 553  
|||||  
Db AAGGATCTGAGAGCCCAAAATCTGCTCAAGAGAGAAATTTCTTCCCAACAGAG 545  
486 AAGGATCTGAGAGCCCAAAATCTGCTCAAGAGAGAAATTTCTTCCCAACAGAG 545  
554 CCCGACACAGAGATCACCTTACATCAAAAGGCTGCTCACTCAACAGAGAGCTGCC 613  
|||||  
Db CCCGACACAGAGATCACCTTACATCAAAAGGCTGCTCACTCAACAGAGAGCTGCC 604  
546 CCCGACACAGAGATCACCTTACATCAAAAGGCTGCTCACTCAACAGAGAGCTGCC 604  
614 TCCACAGAGAGAGCTGGGCTAGAGAAAGAAATCTATTAACCTACAGAGAGCAGATTGAGA 673  
|||||  
Db TCCACAGAGAGAGCTGGGCTAGAGAAAGAAATCTATTAACCTACAGAGAGCAGATTGAGA 663  
605 -TCCACAGAGAGAGCTGGGCTAGAGAAAGAAATCTATTAACCTACAGAGAGCAGATTGAGA 663  
674 CAAAGCATGTAGCCAGAGCCAGGCTGGGCCA-AGAGAGCCAGCTCCAGCTCAACAGAG 732  
|||||  
Db CAAAGCATGTAGCCAGAGCCAGGCTGGGCCAAGAGAGCCAGCTCCAGCTCAACAGAG 723  
664 CAAAGCATGTAGCCAGAGCCAGGCTGGGCCAAGAGAGCCAGCTCCAGCTCAACAGAG 723  
733 AGCTGATCAACACAGCTGCGGCCAGGCTAAACCTTGAGAGCCAAAGAGAGAGCTGCC 790  
|||||  
Db AGCTGATCAACACAGCTGCGGCCAGGCTAAACCTTGAGAGCCAAAGAGAGAGCTGCC 781

RESULT 13  
AI337272 767 bp mRNA linear EST 18-MAR-1999  
LOCUS tb96h07.x1 NCI-CGAP Col6 Homo sapiens cDNA IMAGE:2062237 3'  
DEFINITION similar to gb:L11706 cds1 HORMONE SENSITIVE LIPASE (HUMAN);, mRNA  
sequence.

ACCESSION AI337272  
VERSION AI337272.1 GI:4074199  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. (bases 1 to 767)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Insert length: 3032 std Error: 0.00  
Seq primer: -400p from Gibco  
High quality sequence stop: 502.  
Location/Qualifiers

## FEATURES

1..767  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2062237"  
/clone\_11b="NCI-CGAP\_Col6"  
/tissue\_type="colon tumor, RER+  
/lab\_host="DH10B"  
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP\_Col6 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1057416-1061255, and 1144584-1145351).  
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 134 a 273 c 217 g 138 t 5 others  
ORIGIN

Query Match 18.3%; Score 697; DB 9; Length 767;  
Best Local Similarity 96.4%; Pred. No. 1.3e-147;  
Matches 742; Conservative 0; Mismatches 25; Indels 3; Gaps 3;

1815 GCGCGCGCTTTGGCATGACCCCGAGCTCGTGGGCTGAGTTGAGCGGATCACACAGA 1874  
|||||  
Db 1 GCGCGCGCTTTGGCATGACCCCGAGCTCGTGGGCTGAGTTGAGCGGATCACACAGA 60  
1875 ACCTGAGCTGCACCTTCTGGAAGCTTCTGGAACATCACCGAGATGGAAGTCTATGCT 1934  
|||||  
Db ACCTGAGCTGCACCTTCTGGAAGCTTCTGGAACATCACCGAGATGGAAGTCTATGCT 120  
61 ACCTGAGCTGCACCTTCTGGAAGCTTCTGGAACATCACCGAGATGGAAGTCTATGCT 120  
1935 CTCTGGCAACATGGCATGGGCGACCGATGAGGGTAAAGCGGCTGCTCAGCTGCCACCG 1994  
|||||  
Db CTCTGGCAACATGGCATGGGCGACCGATGAGGGTAAAGCGGCTGCTCAGCTGCCACCG 180  
121 CTCTGGCAACATGGCATGGGCGACCGATGAGGGTAAAGCGGCTGCTCAGCTGCCACCG 180  
1995 AAGCCTTTGAGATGCCATGACTGCGGACCCGACGCTCAAGGCTCAACATCTACCCCCAC 2054  
|||||  
Db AAGCCTTTGAGATGCCATGACTGCGGACCCGACGCTCAAGGCTCAACATCTACCCCCAC 240  
181 AAGCCTTTGAGATGCCATGACTGCGGACCCGACGCTCAAGGCTCAACATCTACCCCCAC 240  
2055 TGGCCACACAGAGCCCTGGGCGCGTCTGTGCAAGCTCATCTCTATGACCTGCGTGAAG 2114  
|||||  
Db TGGCCACACAGAGCCCTGGGCGCGTCTGTGCAAGCTCATCTCTATGACCTGCGTGAAG 300  
241 TGGCCACACAGAGCCCTGGGCGCGTCTGTGCAAGCTCATCTCTATGACCTGCGTGAAG 300  
2115 GACAGGACGTGAGGAGCTCAGAGCTGATTAAGTCAACGAGCCCAAGGAGAGCTGGAGC 2174  
|||||  
Db GACAGGACGTGAGGAGCTCAGAGCTGATTAAGTCAACGAGCCCAAGGAGAGCTGGAGC 360



OY	2175	TTGTGGCCGGGGCCCCAGCAGGAGACCCGGCTGGCGTCCCTGATATAGGACATTCCACGGGG	2234
Db	361	TGTGGCCGGGGCCCCAGCAGGAGACCCGGCTGGCGTCCCTGATATAGGACATTCCACGGGG	420
OY	2235	GTGGCTTTTGTGGCCAGACCTCCAGATCCACAGAGCCTTACTCTAAGAGACTGGGCCAAG	2294
Db	421	GTGGCTTTTGTGGCCAGACCTTACAGATCCACAGAGCCTTACTCTAAGAGACTGGGCCAAG	480
OY	2295	AGCTGGGGCCCCCATCATCTCCATCGACTACTCCCTGGCCCCGTGAGGCCCTTCCCCC	2354
Db	481	AGCTGGGGCCCCCATCATCTCCATCGACTACTCCCTGGCCCCGTGAGGCCCTTCCCCC	539
OY	2355	GTGGCGTGGAGAGTGCCTCTTCCGCTCACTGCTGGGCCATCAAGCACTGCCCTCCTTG	2414
Db	540	GTGGCGTGGAGAGTGCCTCTTCCGCTCACTGCTGGGCCATCAAGCACTGCCCTCCTTG	599
OY	2415	GCTCAACAGGGAGACGAATCTGCTTGGCGGGAGACAGTTCACAGCCGGGAACCTCTGCTTCA	2474
Db	600	GCTCAAC -GGGGAACGAATCTGCTTGGC -GGGAGACAGTTCACAGCCGGGAACCTCTGCTTCA	657
OY	2475	CCGTGGCTTTTGGGCGAGCAGCCTACGGGGGTGGGGTGGCAGATGGCATATGGACGCT	2534
Db	658	NCGTGGCTTTTGGGCGAGCAGCCTCGNNGTGGCGGTGGCCAGATGGCATATGGCAGGAGC	717
OY	2535	ACCGGGCCACATGCTGCAGCCTGGCCGCTCTCCCTCCCGCCTGCTGAGC	2584
Db	718	TACCCGGGCATGCTGCAGCCTGGCCGCTCTTCTCTCCCGCTGCTGAGC	767

RESULT 14		
BC029301	1367 bp	mRNA
LOCUS		linear
BC029301		
DEFINITION	Homo sapiens, clone IMAGE:5169931, mRNA.	
ACCSSION	U00002.0	
DATE		HTC 01-MAY-2002

VERSION	BC029301.1	GI:20381052
KEYWORDS	HTC.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 1367)  
Strausberg, R.  
Direct Submission  
Submitted (01-MAY-2002) National Institutes of Health, Mammalian

REMARK	COMMENT
	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk
	USA National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLUM)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunarane, P. H., Garcia, A. M., Lu, X., Huylk, S. W., Hale, S. M., Yoon, V. S., Kowis, C. R., Lawrence, S., Martin, R. G., Muzny, D. M., Richards, S., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 51 Row: c Column: 24  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis  
 This clone has the following problem: no polyA-tail.

FEATURES	location/Qualifiers
source	1. 1367
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

```

/clone="IMAGE:5169931"
/tissue_type="Brain, adult medulla"
/clone_lib="NTH.MGC.119"
/lab_host="DH10B"
/note="vector: pcMV-SpOrt6"
BASE COUNT      240 a      445 c      453 g      229 t
ORIGIN

```

Query Match	17.6%	Score 668.8	DB 11	Length 1367
Best Local Similarity	99.0%	Pred No. 4.5e-141		
Matches 673	Conservative 0	Mismatches 7	Indels 0	Gaps 0

QY 1143 GACACTACCAGGATACAGCCCTCAAGSCTATCCACACATGGACCTGGCCACATGACAC 1202

[illegible]

Db 748 AGTCGCTGCTACTCTTGGCGGAGACAACATAGCCCTTCCTCGAGCCAGGCTCTGGGG 807

QY 1263 AAACGGCCACGGCGCTGTACAGCGCGTTTTTGTCCCGCTGTACGGGAGCAGGGCGCTGGG 1322

[illegible]

db 868 AGCGGCCCCGGCCGCTGCTGGGTGTGGCGACACCTTTGACCTGGACCCAGAGACAC 927

QY 1383 CGGCCAACGGGTTACCGCAGCCTAGTGCACACAGCCCCCGCTGCTGCGCCGCAACTTCCTGC 1442

[illegible]

Db 988 ACAATCCCGTATGTGGCTCCACCGCCGAGCATCTTCTTCCGACCACGCACAACC 1047

QY 1503 TGGCCGAGCTGGAGGGCCCTACCTGGCTGCCCCCTCACCACAGCTCCGGGCTCTGGTCTACTACG 1562

DD 1040 199CCGAC199GAB99CC1AC199C199CCCC1AC199CCAC199CCG9C199G1C1AC1AC9 110/

Db 1108 CCGAGCCCTGCTGTTACCAATCGCCGGGGGTACTCTTCTTTGAGGGCGACGAGGGGC 1167

QY 1623 TCACCGCCGACTTCCTCCGGGAGTATGTCACGCTGCATTAAGGATGCTTCTATGGCCGCT 1682

Db 1168 TCACCGCCGACTTCCITCCGGGAGTATGTCACGCTGCATAGGGATGCTTCTATGCCCCCT 122/

Db 1228 GCGCTGGGCTTCAGCTTCACGCGCTGCTCATTCGCGGAGAGACCATTCTCCATTGGGCG 1287

1743 TGGTGTCCCTGGGGGAGCACTACAAACGCAACGAGACAGGCTCAGTGTGGCGCCAGCT 1802

Db 1288 TGGTGTCTTCGCGGGAGACACTACAAACGCAACGAGACAGGCTCAGTGTGGCCGCCAGCT 1347

№ п/п	Итого	1257
1	1257	1257
2	1257	1257
3	1257	1257
4	1257	1257
5	1257	1257
6	1257	1257
7	1257	1257
8	1257	1257
9	1257	1257
10	1257	1257
11	1257	1257
12	1257	1257
13	1257	1257
14	1257	1257
15	1257	1257
16	1257	1257
17	1257	1257
18	1257	1257
19	1257	1257
20	1257	1257
21	1257	1257
22	1257	1257
23	1257	1257
24	1257	1257
25	1257	1257
26	1257	1257
27	1257	1257
28	1257	1257
29	1257	1257
30	1257	1257
31	1257	1257
32	1257	1257
33	1257	1257
34	1257	1257
35	1257	1257
36	1257	1257
37	1257	1257
38	1257	1257
39	1257	1257
40	1257	1257
41	1257	1257
42	1257	1257
43	1257	1257
44	1257	1257
45	1257	1257
46	1257	1257
47	1257	1257
48	1257	1257
49	1257	1257
50	1257	1257
51	1257	1257
52	1257	1257
53	1257	1257
54	1257	1257
55	1257	1257
56	1257	1257
57	1257	1257
58	1257	1257
59	1257	1257
60	1257	1257
61	1257	1257
62	1257	1257
63	1257	1257
64	1257	1257
65	1257	1257
66	1257	1257
67	1257	1257
68	1257	1257
69	1257	1257
70	1257	1257
71	1257	1257
72	1257	1257
73	1257	1257
74	1257	1257
75	1257	1257
76	1257	1257
77	1257	1257
78	1257	1257
79	1257	1257
80	1257	1257
81	1257	1257
82	1257	1257
83	1257	1257
84	1257	1257
85	1257	1257
86	1257	1257
87	1257	1257
88	1257	1257
89	1257	1257
90	1257	1257
91	1257	1257
92	1257	1257
93	1257	1257
94	1257	1257
95	1257	1257
96	1257	1257
97	1257	1257
98	1257	1257
99	1257	1257
100	1257	1257

BI826036	LOCUS	663 bp	mRNA	linear	EST 04-OCT-2001
BI826035	LOCUS	663 bp	mRNA	linear	EST 04-OCT-2001

000507074111 mir\_m00\_112 homo sapiens cdna chr10 chr10:210017 2 /

VERSION B1826036.1 GI:15937586  
KEYWORDS EST.

ORGANISM Homo sapiens  
Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

